

CORRIGÉS DES TRAVAUX DIRIGÉS DE L'UE MNB

Mécanique 3A

MÉTHODES NUMÉRIQUES DE BASE

2022-2023, Automne

Jérôme Bastien

Document compilé le 9 août 2023

Le lien original de ce document est le suivant :
<http://utbmjb.chez-alice.fr/Polytech/MNB/TDcorMNB.pdf>

Liste des Travaux Dirigés

Avant-propos	iii
Correction du Travaux Dirigés 1. Interpolation	1
Exercices facultatifs	15
Correction du Travaux Dirigés 2. Intégration	23
Exercices facultatifs	26
Correction du Travaux Dirigés 3. Équations non linéaires	33
Correction du Travaux Dirigés 4. Équations différentielles	51
Correction du Travaux Dirigés 5. Systèmes d'équations linéaires	59
Exercices facultatifs	62
Annexe A. Rappels théoriques	63
A.1. Interpolation	63
A.2. Intégration	68
A.3. Équations non-linéaires	68
A.4. Équations différentielles	73
Annexe B. Simulations numériques sur l'erreur d'interpolation	75
Annexe C. Étude du majorant exact d'une dérivée quatrième	83
Annexe D. Définition et utilisation de la fonction W de Lambert	87
D.1. Définition de la fonction W de Lambert	87
D.2. Utilisation de la fonction W de Lambert : résolution de l'équation $ae^x + bx + c = 0$	88
D.3. Utilisation de la fonction W de Lambert : résolution de l'équation $x^x = z$	90
Annexe E. Méthode de dichotomie ou de bisection (sous la forme d'un exercice corrigé)	91
Énoncé	91
Corrigé	92
Annexe F. Convergence globale de la méthode du point fixe	93
F.1. Cas particuliers	93
F.2. Cas général	93
Annexe G. Étude de la convergence globale de la suite définie par $u_0 \in \mathbb{R}_+^*$ et $u_{n+1} = 1/2(u_n + A/u_n)$, sous la forme d'un	
Énoncé	95
Corrigé	95
Annexe H. Étude du zéro de la fonction $e^x + x - K$, sous la forme d'un problème corrigé	101
Énoncé	101

Corrigé	104
Annexe I. Étude et calcul de l tel que $l = \cos l$ sous la forme d'un problème corrigé	125
Énoncé	125
Corrigé	126
Annexe J. Dégénérescence de la méthode de Newton et méthode de Newton modifiée	135
Annexe K. Racines multiples	141
K.1. Racines multiples d'un polynôme	141
K.2. Racines multiples d'une fonction quelconque	143
Annexe L. Un théorème de point fixe	145
Annexe M. Convergence des méthodes d'ordre p	147
Annexe N. Compléments sur la divergence de la méthode du point fixe (sous forme d'un exercice corrigé)	151
Énoncé	151
Corrigé	151
Quelques remarques	158
Bibliographie	163

Avant-propos

Ce polycopié constitue les corrigés de TD de Méthodes Numériques de Base du département Mécanique 3A (2022-2023, Automne).

Un certain nombre d'exercices sont extraits de [BM03]; certaines corrections ne sont pas données, mais elles figurent dans cet ouvrage, disponible à la BU.

Les résultats théoriques, issus du cours et utilisés dans ces corrections sont tous rappelés dans l'annexe A.

Des notes en petits caractères comme suit pourront être omises en première lecture :

Attention, passage difficile! \diamond

Ce polycopié de corrigé de TD est normalement disponible à la fois

- en ligne sur <http://utbmjb.chez-alice.fr/Polytech/index.html> à la rubrique habituelle ;
- en cas de problème internet, sur le réseau de l'université Lyon I : il faut aller sur :
 - 'Poste de travail',
 - puis sur le répertoire 'P:' (appelé aussi '\\teraetu\Enseignants'),
 - puis 'jerome.bastien',
 - puis 'Polytech',
 - puis 'Mécanique 3A'.
 - enfin sur 'MNB'.

Interpolation

Pour tous les calculs de polynôme d'interpolation par la méthode de Newton, on consultera la définition A.3 page 63 de l'annexe A.

CORRECTION DE L'EXERCICE 1.1.

(1) Ici, on a $n = 1$.

Chacun des polynômes de Lagrange l_i (de degré 1) est donné par la formule :

$$\forall i \in \{0, \dots, n\}, \quad l_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j}. \quad (1.1)$$

On a donc successivement

$$l_0(x) = \frac{(x - 5)}{(3 - 5)},$$

$$l_1(x) = \frac{(x - 3)}{(5 - 3)}.$$

soit encore après calculs :

$$l_0(x) = -1/2 x + 5/2, \quad (1.2a)$$

$$l_1(x) = 1/2 x - 3/2. \quad (1.2b)$$

Ensuite, le polynôme interpolateur de degré 1, $\Pi_n(g)$, est donné par la formule :

$$\Pi_n(g)(x) = \sum_{i=0}^n g(x_i) l_i(x). \quad (1.3)$$

Ici, on a donc :

$$\Pi_n(g)(x) = g(x_0) l_0(x) + g(x_1) l_1(x).$$

Après calculs, il vient :

$$\Pi_n(g)(x) = 2x - 8. \quad (1.4)$$

Pour $\alpha = 4.8$, on obtient alors :

$$\Pi_n(g)(\alpha) = 1.6, \quad (1.5)$$

ce qui constitue une valeur approchée de $g(\alpha)$.

Sur la figure 1.1, ont été tracés les polynômes de Lagrange l_0 et l_1 et le polynôme interpolateur.

(2) Ici, on a $n = 2$.

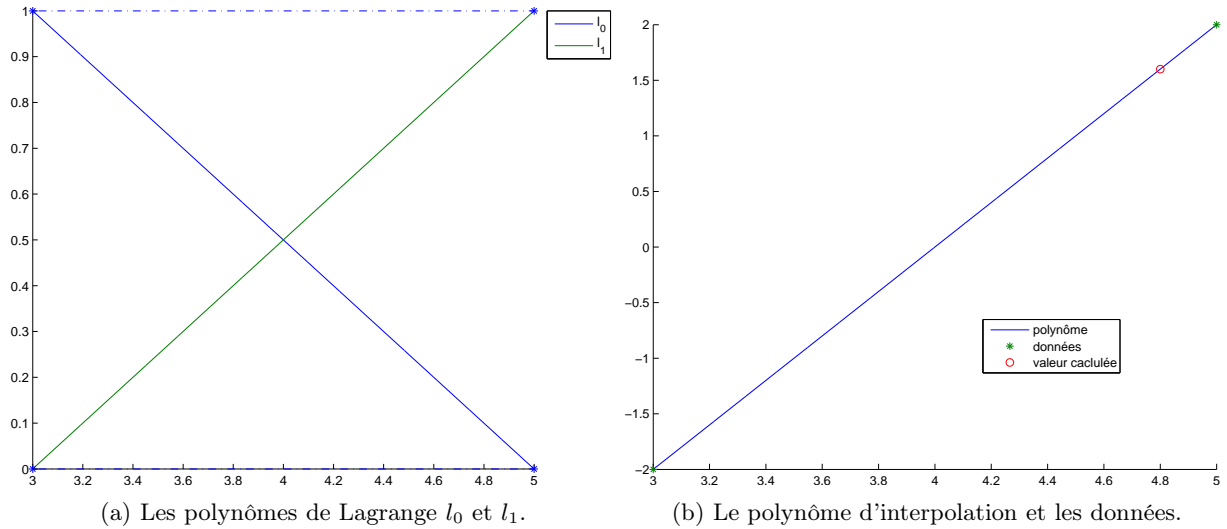


FIGURE 1.1. L'interpolation de Lagrange.

Chacun des polynômes de Lagrange l_i (de degré 2) est donné par la formule (1.1). On a donc successivement

$$l_0(x) = \frac{(x-5)(x-8)}{(3-5)(3-8)},$$

$$l_1(x) = \frac{(x-3)(x-8)}{(5-3)(5-8)},$$

$$l_2(x) = \frac{(x-3)(x-5)}{(8-3)(8-5)}.$$

soit encore après calculs :

$$l_0(x) = 1/10 x^2 - \frac{13}{10} x + 4, \quad (1.6a)$$

$$l_1(x) = -1/6 x^2 + \frac{11}{6} x - 4, \quad (1.6b)$$

$$l_2(x) = 1/15 x^2 - \frac{8}{15} x + 1. \quad (1.6c)$$

Ensuite, le polynôme interpolateur de degré 2, $\Pi_n(g)$, est donné par la formule (1.3). Ici, on a donc :

$$\Pi_n(g)(x) = g(x_0)l_0(x) + g(x_1)l_1(x) + g(x_2)l_2(x).$$

Après calculs, il vient :

$$\Pi_n(g)(x) = -1/3 x^2 + 14/3 x - 13. \quad (1.7)$$

Pour $\alpha = 4.8$, on obtient alors :

$$\Pi_n(g)(\alpha) = 1.72000000000000000000000000000001, \quad (1.8)$$

ce qui constitue une valeur approchée de $g(\alpha)$.

Sur la figure 1.2, ont été tracés les polynômes de Lagrange l_0 , l_1 et l_2 et le polynôme interpolateur.

(3) (a) Ici, on a $n = 1$.

Pour calculer le polynôme sous la forme de Newton, on détermine tout d'abord les différences divisées $g[x_i, \dots, x_{i+k}]$ données dans le tableau 1.1. Ensuite, on n'utilise plus que les différences divisées qui

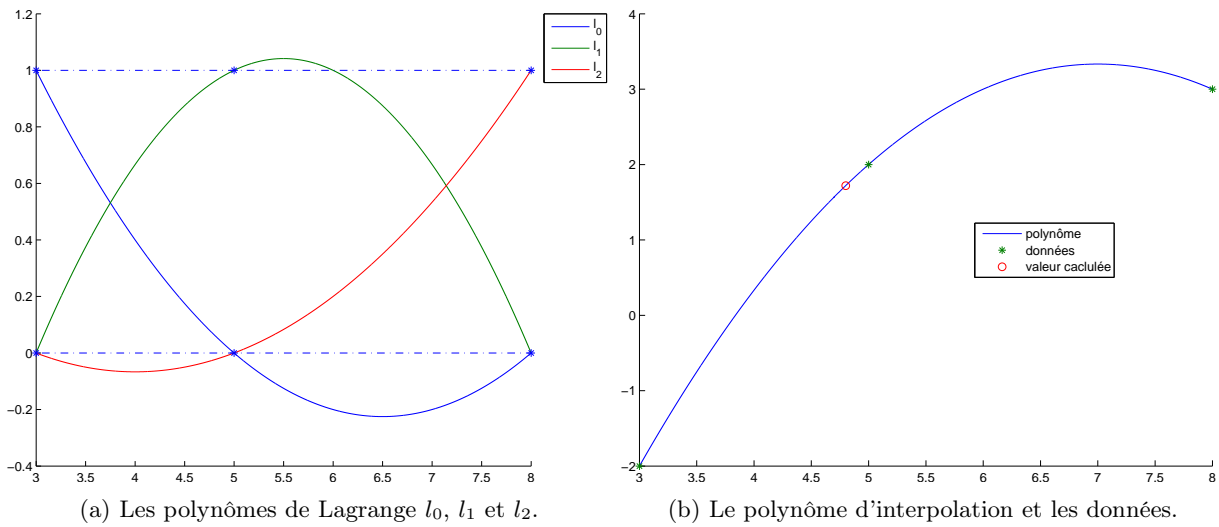


FIGURE 1.2. L'interpolation de Lagrange.

$x_i \setminus k$	0	1
$x_0 = 3$	-2	
$x_1 = 5$	2	2

TABLE 1.1. Différences divisées de g .

sont encadrées et le polynôme interpolateur de degré 1, $\Pi_n(g)$, est donné par la formule :

$$\Pi_n(g)(x) = \sum_{i=0}^n g[x_0, \dots, x_i](x - x_0) \dots (x - x_{i-1}). \quad (1.9)$$

Ici, on a donc :

$$\Pi_n(g)(x) = g[x_0] + g[x_0, x_1](x - x_0).$$

On a

$$x - x_0 = x - 3,$$

Après calculs, il vient :

$$\Pi_n(g)(x) = 2x - 8. \quad (1.10)$$

Pour $\alpha = 4.8$, on retrouve alors la valeur donnée par (1.5).

(b)

Ici, on a $n = 2$. On n'est pas obligé de reprendre tous les calculs. Il suffit de rajouter le point x_2 dans le tableau des différences divisées et la valeur correspondante $g(x_2)$, comme le montre le tableau 1.2. Ensuite, on utilise la formule (1.9), pour passer de $n - 1$ à n ; il suffit de rajouter le polynôme $g[x_0, \dots, x_n](x - x_0) \dots (x - x_{n-1})$ au polynôme $\Pi_{n-1}(g)$ pour obtenir le polynôme $\Pi_n(g)$. On obtient le polynôme ci-dessus.

$x_i \setminus k$	0	1	2
$x_0 = 3$	-2		
$x_1 = 5$	2	2	-1/3
$x_2 = 8$	3	1/3	

TABLE 1.2. Différences divisées de g .

Pour $\alpha = 4.8$, on retrouve alors la valeur donnée par (1.8).

- (4) La seconde solution est plus rapide, car fait appel aux calculs précédents, ce qui était le but poursuivi par « l'inventeur des différences divisées », Isaac Newton. On pourra consulter par exemple <http://www.unige.ch/~wanner/teaching/Numi/Numi2.pdf>

CORRECTION DE L'EXERCICE 1.2.

- (1) (a) Déterminons tout d'abord les points de Chebytsheff.

On peut procéder de deux façons.

- (i) Les formules (A.19) donnent numériquement

$$x_0 = 5, \quad (1.11a)$$

$$x_1 = 5.25, \quad (1.11b)$$

$$x_2 = 5.75, \quad (1.11c)$$

$$x_3 = 6. \quad (1.11d)$$

- (ii) On peut aussi utiliser l'interprétation de la remarque A.9, avec $a = 5$ et $b = 6$. On considère le demi-cercle de centre d'abscisse

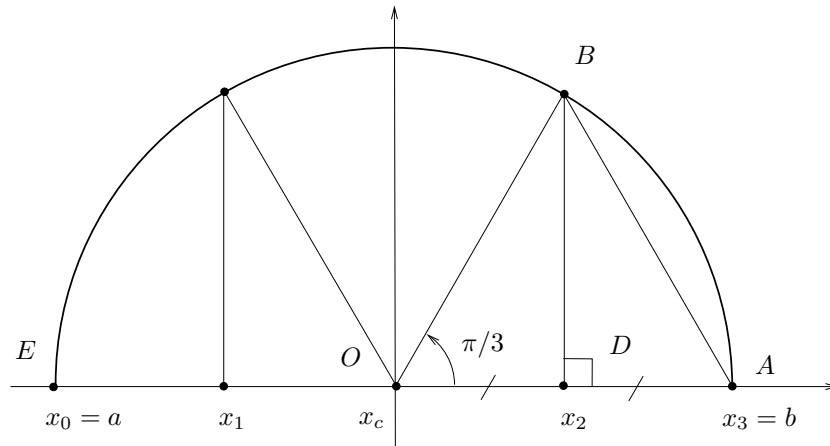
$$x_c = \frac{a+b}{2} = 5.50 \quad (1.12)$$

et de rayon

$$R = \frac{b-a}{2} = 0.50, \quad (1.13)$$

c'est-à-dire passant par les points d'abscisses a et b . On divise ce demi-cercle en en $n = 3$ parties égales; on a donc 4 points définis par les angles $i\pi/3$ pour $0 \leq i \leq 3$. Voir la figure 1.3 page suivante. Les $n+1 = 4$ abscisses de ces points ne sont autres que les points définis par (A.19) dans le cas $n = 3$. Il est clair que x_0 et x_3 sont égaux respectivement à a et b et on retrouve (1.11a) et (1.11d). Géométriquement, on constate que le triangle OAB est un triangle équilatéral (deux coté égaux et l'angle au sommet égal à $\pi/3$). Ainsi, le point D , le pied de la hauteur (BD) est le milieu de $[OA]$ et donc $OD = R/2$ et donc $ED = 3/4 R$. Autrement dit,

$$x_2 = a + \frac{3}{2}R$$

FIGURE 1.3. Les points de Tchebycheff pour $n = 3$.

et donc, grâce à (1.13), il vient

$$\begin{aligned} x_2 &= a + \frac{3}{4}(b - a), \\ &= \frac{1}{4}a + \frac{3}{4}b, \\ &= 5.75, \end{aligned}$$

et on retrouve (1.11c). On fait de même pour (1.11d).

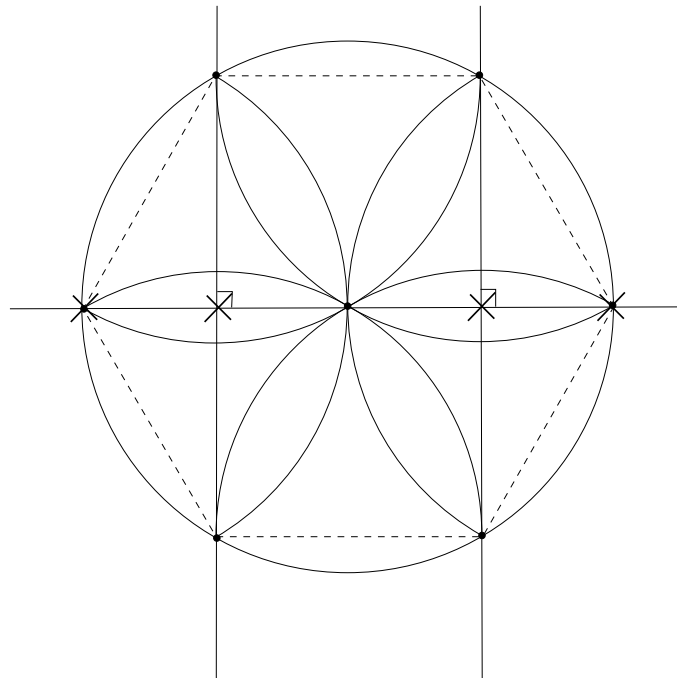


FIGURE 1.4. La rosace (et l'hexagone qui la sous-tend, en pointillé), avec les quatre points de Tchebycheff, repérés par quatre croix

On peut aussi faire cette construction à la règle et au compas, puisque ce cas particulier est fondé sur la construction d'un triangle équilatéral. Voir par exemple la construction de l'hexagone régulier, issu de [Bas22, Section "Quelques exercices", dans l'annexe "] et la figure 1.4 page précédente, adaptée de la référence citée, qui montre la jolie rosace habituelle.

(b) Déterminons maintenant le polynôme d'interpolation.

$x_i \setminus k$	0	1	2	3
$x_0 = 5$	2.564949			
$x_1 = 5.250000$	2.602690	0.150961	-0.010725	
$x_2 = 5.750000$	2.674149	0.142918	-0.009749	0.000976
$x_3 = 6$	2.708050	0.135606		

TABLE 1.3. Différences divisées de f .

Pour calculer le polynôme sous la forme de Newton, on détermine tout d'abord les différences divisées $f[x_i, \dots, x_{i+k}]$ données dans le tableau 1.3. Ensuite, on n'utilise plus que les différences divisées qui sont encadrées et le polynôme interpolateur de degré 3, $\Pi_3(f)$, est donné par la formule :

$$\Pi_3(f)(x) = \sum_{i=0}^n f[x_0, \dots, x_i](x - x_0) \dots (x - x_{i-1}). \quad (1.14)$$

Ici, on a donc :

$$\Pi_3(f)(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) + f[x_0, x_1, x_2, x_3](x - x_0)(x - x_1)(x - x_2).$$

On a successivement

$$\begin{aligned} x - x_0 &= x - 5, \\ (x - x_0)(x - x_1) &= x^2 - 10.250000x + 26.250000, \\ (x - x_0)(x - x_1)(x - x_2) &= x^3 - 16x^2 + 85.187500x - 150.937500. \end{aligned}$$

Après calculs, il vient :

$$\Pi_3(f)(x) = 0.000976x^3 - 0.026333x^2 + 0.343992x + 1.381377. \quad (1.15)$$

Voir la figure 1.5 page suivante.

(2) On utilise le théorème A.4 : il existe ξ tel que

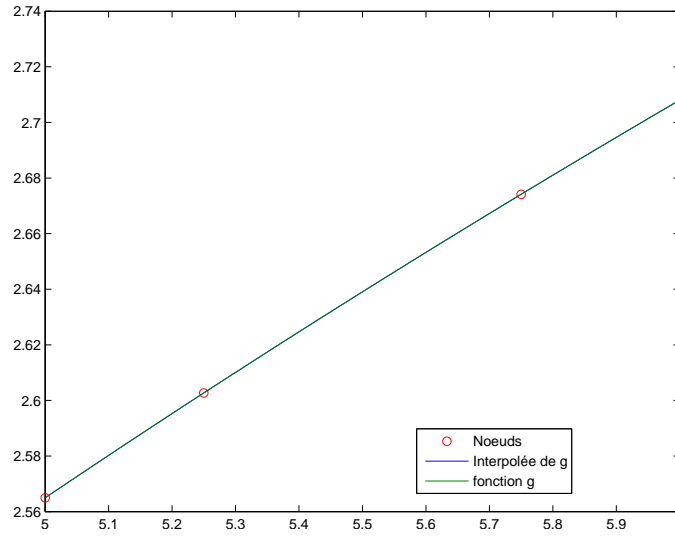
$$E_3(f)(x) = \frac{f^{(4)}(\xi)}{4!} \omega_4(x),$$

et donc

$$|E_3(f)| \leq \frac{\max_{\xi} |f^{(4)}(\xi)|}{4!} \max_x |\omega_4(x)|,$$

Graphiquement, on constate que

$$\max(|\omega_4|) \approx 0.016000.$$

FIGURE 1.5. La fonction g et son interpolée.

Par ailleurs, on peut vérifier que

$$f^{(4)}(x) = -96 (2x + 3)^{-4}, \quad (1.16)$$

et puisque sur $[a, b]$, $x \geq 5$, on a

$$|f^{(4)}(x)| \leq 3.36122 \cdot 10^{-3}, \quad (1.17)$$

dont on déduit que

$$|E_3(f)| \leq 1.60000 \cdot 10^{-2} \times \frac{1}{4!} \times 3.36122 \cdot 10^{-3} = 2.24081 \cdot 10^{-6}. \quad (1.18)$$

REMARQUE 1.1. Si utilise la fonction `maxabsfun`, fournie sur le site habituel, pour déterminer avec plus de soin, un majorant exact de $f^{(4)}$, on obtient

$$M = 3.36122 \cdot 10^{-3}.$$

qui est identique à (1.17).

Voir figure 1.6.

Dans ce cas, l'estimation (1.18) est identique puisque l'on obtient

$$|E_3(f)| \leq 1.60000 \cdot 10^{-2} \times \frac{1}{4!} \times 3.36122 \cdot 10^{-3} = 2.24081 \cdot 10^{-6}.$$

CORRECTION DE L'EXERCICE 1.3.

(1) D'après le théorème A.7, on a, pour des nœuds équirépartis de l'intervalle $[a, b]$

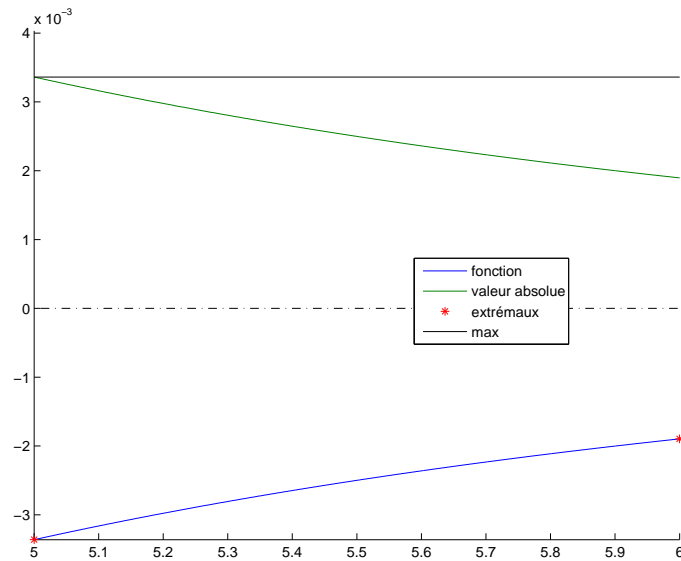
$$E_n(f) \leq \frac{1}{4(n+1)} \left(\frac{b-a}{n} \right)^{n+1} \max_{x \in [a, b]} |f^{(n+1)}(x)|. \quad (1.19)$$

La dérivée n -ième de f est majorée, en valeur absolue, par $1/3^n$. D'où

$$E_n(f) \leq \frac{1}{4(n+1)n^{n+1}3^{n+1}},$$

soit encore

$$E_n(f) \leq \frac{1}{4(n+1)(3n)^{n+1}}. \quad (1.20)$$

FIGURE 1.6. Extréma de la fonction f définie par (1.16).

(2) On sait que cette quantité tend vers 0 quand n tend vers l'infini. Nous proposons deux méthodes légèrement différentes.

(a) De façon brutale, on détermine à la main, le plus petit n tel que cette quantité soit inférieure à $\varepsilon = 10^{-4}$. C'est-à-dire, on teste successivement $n = 1, n = 2, \dots$. On trouve

$$n = 3. \quad (1.21)$$

(b) De façon plus subtile, remarquons que l'on a successivement les majorations suivantes : pour tout n entier, on a

$$\begin{aligned} 4(n+1)(3n)^{n+1} &\geq 4(3n)^{n+1}, \\ &= 4 \times 3^{n+1} n^{n+1}, \\ &= 12 \times 3^n n^{n+1}, \\ &\geq 12 \times 3^n n^n, \\ &= 12 \times (3n)^n, \end{aligned}$$

et donc l'inégalité

$$\frac{1}{4(n+1)(3n)^{n+1}} \leq \varepsilon \quad (1.22)$$

est vraie si

$$\frac{1}{12(3n)^n} \leq \varepsilon, \quad (1.23)$$

ce qui est équivalent à

$$(3n)^n \geq \frac{1}{12\varepsilon},$$

et donc à

$$(3n)^{3n} \geq \frac{1}{(12\varepsilon)^3}.$$

Par croissance, il suffit donc de trouver le plus petit entier n_0 tel que

$$(3n_0)^{3n_0} = \frac{1}{(12\varepsilon)^3}.$$

D'après la section D.3 de l'annexe D page 87, cela est équivalent à

$$x = \frac{\ln z}{W(\ln z)},$$

où

$$x = 3n_0 \text{ et } z = \frac{1}{(12\varepsilon)^3}.$$

On trouve donc, dans \mathbb{R} :

$$n_0 = \frac{\ln z}{3W(\ln z)},$$

soit numériquement :

$$n_0 = 3.0417459.$$

et donc, dans \mathbb{N} :

$$n_0 = 4,$$

ce qui est un peu plus grand que la valeur donnée par (1.21), ce qui est normal car on a raisonné en condition suffisante, en remplaçant l'inégalité (1.22) par l'inégalité (1.23) moins forte. Au final, cette méthode est plus pessimiste !

(3) Tous les résultats sont identiques pour la fonction g .

REMARQUE 1.2. Dans cet exercice, on a montré que l'erreur d'interpolation entre f et son polynôme d'interpolation p_n , défini par $n + 1$ points équirépartis est donnée par (1.19) dans le cas général et par (1.20) dans le cas particulier étudié.

Ainsi, cette erreur tend vers zéro quand n tend vers l'infini. Attention, ce résultat n'est vrai qu'en théorie ; en pratique, du fait des arrondis de calculs, l'erreur d'interpolation entre f et p_n ne tend pas vers zéro quand n grandit.

De plus, dans [BM03, Exercice 2.3 p. 53 et 239 et TP 2.D p. 65], nous avons mis en avant le fait que l'on peut évaluer numériquement le polynôme p_n de trois façons :

(1) en utilisant la forme de Newton donnée par

$$p_n(x) = \sum_{i=0}^n f[x_0, \dots, x_i](x - x_0)\dots(x - x_{i-1}), \quad (1.24)$$

et l'algorithme d'évaluation de Horner ¹

1. qui consiste en fait à évaluer $p_n(x)$ de la façon suivante : si on note $a_0 = f[x_0]$, $a_1 = f[x_0, x_1]$, $a_2 = f[x_0, x_1, x_2], \dots$, $a_n = f[x_0, \dots, x_n]$, alors on calcule p_n défini par (1.24) de la façon suivante (dans cet ordre, sans exprimer ce qui suit dans la base canonique) :

$$p_1(x) = a_0 + (x - x_0)a_1, \quad (1.25a)$$

si $n = 1$,

$$p_2(x) = a_0 + (x - x_0)(a_1 + (x - x_1)a_2), \quad (1.25b)$$

si $n = 2$ et de façon plus générale :

$$p_n(x) = a_0 + (x - x_0)(a_1 + (x - x_1)(a_2 + (x - x_2)(a_3 + (x - x_3)(\dots(a_{n-2} + (x - x_{n-2})(a_{n-1} + (x - x_{n-1})a_n))))), \quad (1.25c)$$

de façon, entre autres, à diminuer le nombre d'opérations à effectuer, notamment le calcul des produits $(x - x_0)\dots(x - x_i)$. (voir [BM03, Algorithme 2.1 d'Horner p. 38 et TP 2.A p. 61])

(2) en utilisant sa forme canonique²

$$p_n(x) = \sum_{i=0}^n b_i x^i ; \quad (1.26)$$

(3) en utilisant son écriture dans la base des polynômes de Lagrange l_i

$$p_n(x) = \sum_{i=0}^n f(x_i) l_i(x), \quad (1.27)$$

où

$$l_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j}. \quad (1.28)$$

L'erreur est intrinséquement liée à l'utilisation d'un calculateur par le cumul d'imprécisions liées aux arrondis. Cependant, les trois méthodes n'ont pas la même sensibilité numérique. Nous donnons en annexe B deux simulations numériques mettant cela en évidence sur les fonctions f et g définies dans l'énoncé.

CORRECTION DE L'EXERCICE 1.4.

D'après le théorème A.4, on a, sur chaque intervalle $[x_i, x_{i+1}] = [ih, (i+1)h]$,

$$E_1(x) = \frac{f^{(2)}(\xi)}{2!} (x - x_i)(x - x_{i+1}).$$

Sur $[0, 1]$, on peut majorer brutalement chacun des produits $(x - x_i)$ par h et donc $|(x - x_i)(x - x_{i+1})|$ par h^2 . On peut aussi être plus précis et montrer que la fonction $x \mapsto (a - x)(b - x)$ a un maximum égal à $\frac{(a-b)^2}{4}$. On majore donc $|(x - x_i)(x - x_{i+1})|$ par $h^2/4$. On a donc

$$|E_1(x)| \leq \max_{\xi \in [0,1]} |f^{(2)}(\xi)| \frac{h^2}{2 \times 4} = \max_{\xi \in [0,1]} |f^{(2)}(\xi)| \frac{h^2}{8}.$$

On peut aussi retrouver cette formule grâce au théorème A.11. Par ailleurs, on a

$$f^{(2)}(x) = 2e^{-x^2}(-1 + 2x^2). \quad (1.29)$$

On peut utiliser brutalement l'inégalité triangulaire

$$\forall a, b \in \mathbb{R}, \quad |a + b| \leq |a| + |b|. \quad (1.30)$$

On écrit donc

$$\forall x \in [0, 1], \quad |f^{(2)}(x)| \leq 2 \left| e^{-x^2} \right| |2x^2 - 1|, \quad (1.31)$$

ce qui donne

$$\forall x \in [0, 1], \quad |f^{(2)}(x)| \leq 2e^{-x^2} |2x^2 - 1|, \quad (1.32)$$

et donc

$$\forall x \in [0, 1], \quad |f^{(2)}(x)| \leq 2 |2x^2 - 1|, \quad (1.33)$$

dont on déduit, grâce à (1.30),

$$\forall x \in [0, 1], \quad |f^{(2)}(x)| \leq 2 (|2x^2| + |-1|), \quad (1.34)$$

et donc

$$\forall x \in [0, 1], \quad |f^{(2)}(x)| \leq 2(2x^2 + 1), \quad (1.35)$$

et enfin, puisque sur $[0, 1]$, $x^2 \leq 1$:

$$\forall x \in [0, 1], \quad |f^{(2)}(x)| \leq 2(2 + 1) = 6, \quad (1.36)$$

2. que l'on obtient grâce à la forme de Newton et par exemple [BM03, Exercice 2.7 p. 56, Algorithme 6.1 p. 246].

et un majorant M de $|f^{(2)}(x)|$ sur $[0, 1]$ est donné par

$$M = 6, \quad (1.37)$$

Bref

$$|E_1(x)| \leq \frac{3}{4}h^2.$$

Ainsi, pour avoir une erreur inférieure à $\varepsilon = 5 \cdot 10^{-7}$, il suffit que

$$\frac{3}{4}h^2 \leq \varepsilon,$$

soit

$$h \leq \sqrt{\frac{4}{3}\varepsilon}, \quad (1.38)$$

soit encore

$$\frac{1}{N} \leq \sqrt{\frac{4}{3}\varepsilon}.$$

On a donc

$$N \geq \sqrt{\frac{3}{4\varepsilon}} = 1224.744871.$$

Puisque N est entier, on choisit

$$N = 1225. \quad (1.39)$$

REMARQUE 1.3. On peut améliorer la valeur de M donnée par (1.37) de trois façons.

- (1) Comme l'a remarqué M. Thioliere en 2018, il suffit de reprendre l'inégalité (1.31). D'une part, on utilise le fait que $x \mapsto 2x^2 - 1$ est croissante sur $[0, 1]$ et qu'elle vaut -1 en 0 et 1 en 1 ; cette fonction est donc comprise entre -1 et 1 et en valeur absolue, elle est donc majorée par 1. On a donc

$$\forall x \in [0, 1], \quad |2x^2 - 1| \leq 1. \quad (1.40)$$

D'autre part, la fonction $x \mapsto e^{-x^2}$ est positive, décroissante sur $[0, 1]$ et elle vaut 1 en 0 ; elle est donc majorée par 1, en valeur absolue. On a donc

$$\forall x \in [0, 1], \quad e^{-x^2} \leq 1. \quad (1.41)$$

Bref, de (1.32) (1.40) et (1.41), on déduit

$$\forall x \in [0, 1], \quad |f^{(2)}(x)| \leq 2 \times 1, \quad (1.42)$$

et donc

$$\forall x \in [0, 1], \quad |f^{(2)}(x)| \leq 2, \quad (1.43)$$

et on remplace donc (1.37) par

$$M = 2. \quad (1.44)$$

Dans ce cas, M est optimal puisque'il est atteint par f :

$$|f^{(2)}(0)| = M. \quad (1.45)$$

- (2) On peut aussi, remarquer que d'après (1.29), on a

$$f^{(3)}(x) = 4xe^{-x^2}(3 - 2x^2). \quad (1.46)$$

La fonction $x \mapsto 3 - 2x^2$ est décroissante sur $[0, 1]$, elle vaut 0 en 1 et 1 en 1 et est donc positive sur $[0, 1]$; Ainsi, $f^{(3)}$ est positive sur $[0, 1]$ et $f^{(2)}$ est donc croissante sur $[0, 1]$. Elle est donc comprise entre $f^{(2)}(0) = -2$ et $f^{(2)}(1) = 2e^{-1}$, tous les deux inférieurs à 2, en valeur absolue. On retrouve donc (1.44).

- (3) On peut aussi, de façon informatique, utiliser la fonction `maxabsfun`, fournie sur le site habituel, pour déterminer avec plus de soin, un majorant exact de $f^{(2)}$, on obtient

$$M = 2.000000000, \quad (1.47)$$

et on retrouve donc (1.44).

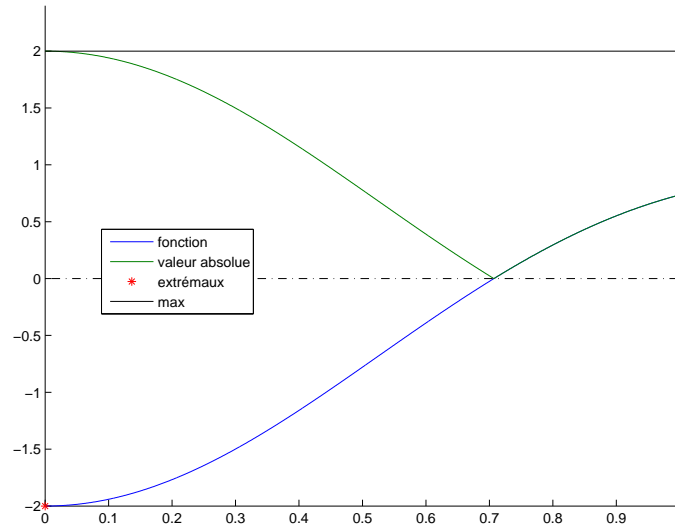


FIGURE 1.7. Extréma de la fonction g définie par (1.29).

Voir figure 1.7.

En utilisant (1.44) ou (1.47), l'estimation (1.38) est remplacée par

$$h \leq \sqrt{\frac{8}{M}} \varepsilon,$$

et donc puisque

$$N = \frac{1}{h},$$

l'estimation (1.39) est remplacée un peu plus avantageusement par

$$N = 708.$$

CORRECTION DE L'EXERCICE 1.5.

- (1) Parmi les différents points t de l'énoncé, on considère quatre points successifs qui encadrent "au mieux" la donnée $\tau = 3.100000$. On utilise tout d'abord le fait que l'on fait de l'"inter"polation (la donnée τ est à l'intérieur des points t_i). De plus, l'écart entre τ et les t_i doit être le plus faible possible. En outre, l'écart entre les différents t_i doit être le plus faible possible. Enfin, on admet que la situation doit être la plus symétrique possible. Finalement, on choisit donc :

$$t_0 = 2.50, \quad (1.48a)$$

$$t_1 = 3, \quad (1.48b)$$

$$t_2 = 3.50, \quad (1.48c)$$

$$t_3 = 4. \quad (1.48d)$$

$t_i \setminus k$	0	1	2	3
$t_0 = 2.500000$	7.829100			
		2.734400		
$t_1 = 3$	9.196300		0.075000	
		2.809400		-0.053867
$t_2 = 3.500000$	10.601000		-0.005800	
		2.803600		
$t_3 = 4$	12.002800			

TABLE 1.4. Différences divisées de f .

Pour calculer le polynôme sous la forme de Newton, on détermine tout d'abord les différences divisées $f[t_i, \dots, t_{i+k}]$ données dans le tableau 1.4. Ensuite, on n'utilise plus que les différences divisées qui sont encadrées et le polynôme interpolateur de degré 3, Π_3 , est donné par la formule :

$$\Pi_3(t) = \sum_{i=0}^n f[t_0, \dots, t_i](t - t_0) \dots (t - t_{i-1}). \quad (1.49)$$

Ici, on a donc :

$$\Pi_3(t) = f[t_0] + f[t_0, t_1](t - t_0) + f[t_0, t_1, t_2](t - t_0)(t - t_1) + f[t_0, t_1, t_2, t_3](t - t_0)(t - t_1)(t - t_2).$$

On a successivement

$$\begin{aligned} t - t_0 &= t - 2.500000, \\ (t - t_0)(t - t_1) &= t^2 - 5.500000t + 7.500000, \\ (t - t_0)(t - t_1)(t - t_2) &= t^3 - 9t^2 + 26.750000t - 26.250000. \end{aligned}$$

Après calculs, il vient :

$$\Pi_3(t) = -0.053867t^3 + 0.559800t^2 + 0.880967t + 2.969600. \quad (1.50)$$

On en déduit

$$\Pi_3(\tau) = \Pi_3(3.100000) = 9.475533.$$

Voir la figure 1.8.

REMARQUE 1.4. Au lieu de prendre (1.48) on considère cette fois-ci les points t_i définis par

$$t_0 = 1, \quad (1.51a)$$

$$t_1 = 1.500000, \quad (1.51b)$$

$$t_2 = 2, \quad (1.51c)$$

$$t_3 = 2.500000, \quad (1.51d)$$

$$t_4 = 3, \quad (1.51e)$$

$$t_5 = 3.500000, \quad (1.51f)$$

$$t_6 = 4. \quad (1.51g)$$

Une autre façon plus complète de procéder et de déterminer tous les polynômes de degrés 3 définis par 4 points qui se suivent parmi les t_i , qui peuvent même ne pas encadrer τ ! On dispose des 4 possibilités

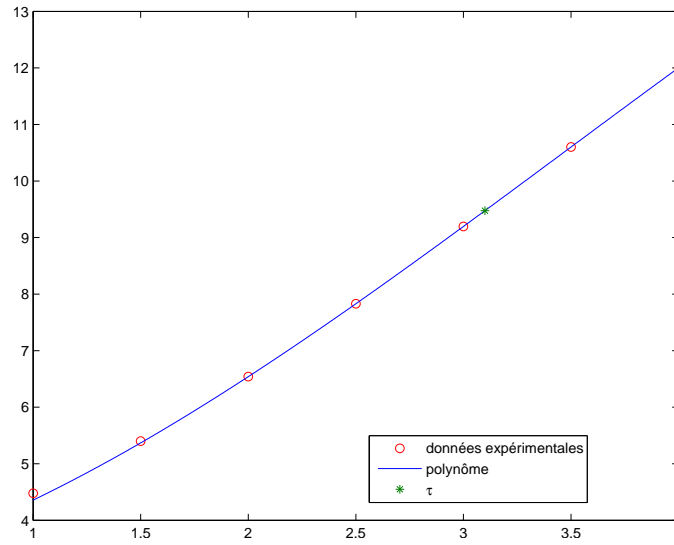


FIGURE 1.8. Les données expérimentales et la courbe de degré 3.

différentes données par $i \in \{1, 2, 3, 4\}$, $i \in \{2, 3, 4, 5\}$, $i \in \{3, 4, 5, 6\}$ et $i \in \{3, 4, 5, 6\}$. On calcule pour ces 4 possibilités les différences divisées puis $\Pi_3(\tau)$. On trouve

$$\Pi_3(\tau) = 9.460207,$$

$$\Pi_3(\tau) = 9.473442,$$

$$\Pi_3(\tau) = 9.475581,$$

$$\Pi_3(\tau) = 9.475533.$$

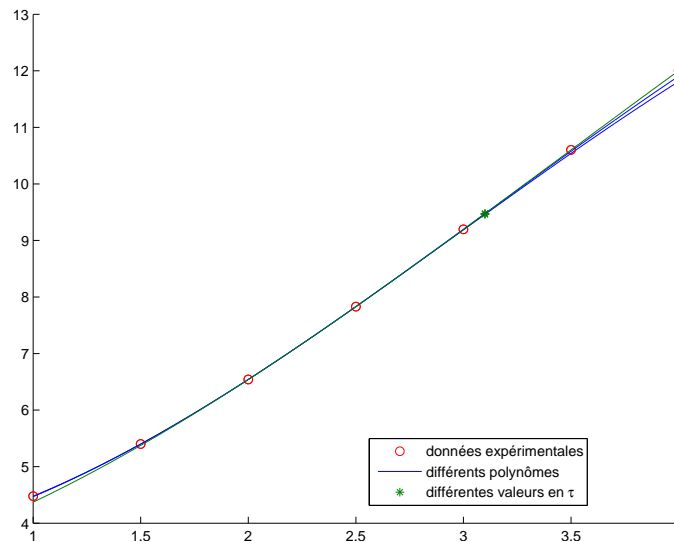


FIGURE 1.9. Les données expérimentales et les différentes courbes de degré 3.

Voir la figure 1.9. On constate que les quatre solutions sont quasi identiques.

(2) L'expression analytique de l'erreur est donnée par

$$E_3(x) = \frac{f^{(4)}(\xi)}{4!} \omega_4(x), \quad (1.52)$$

où ici f et ξ ne sont pas connus. Cette formule provient de la propriété A.4 utilisée ici avec $n = 3$.

(3) Il est donc impossible de calculer une telle erreur. Cependant, on peut supposer par exemple que f est polynomiale de degré 4, ou ce qui revient au même que $f^{(4)}$ ne varie pas. On a alors une approximation de l'erreur, obtenue en prenant un point supplémentaire t_{n+1} :

$$E_3(x) \approx f[t_0, \dots, t_{n+1}](x - t_0) \dots (x - t_n), \quad (1.53)$$

utilisée ici avec $n = 3$. Cela provient de (1.52) ; en effet, si f est polynomiale de degré 4, le coefficient dominant de f est $f[t_0, \dots, t_{n+1}]$ puisque c'est le coefficient dominant de $f = \Pi_4$ sur la forme de Newton. On a alors pour tout ξ , puisque $f^{(4)}$ est constant :

$$f^{(4)}(\xi) = 4! f[t_0, \dots, t_4],$$

ce qui donne, réinjecté dans (1.52) :

$$E_3(x) = f[t_0, \dots, t_4] \omega_4(x).$$

Si cette fois-ci, $f^{(4)}$ varie peu, cela devient une approximation et on a donc, dans tous les cas (1.53). Prenons $t_0 = 2.500000$, $t_1 = 3$, $t_2 = 3$, $t_3 = 4$ et $t_4 = 2$. On complète le tableau déjà fait dans la

$t_i \setminus k$	0	1	2	3	4
$t_0 = 2.500000$	7.829100				
$t_1 = 3$	9.196300	2.734400	0.075000		
$t_2 = 3.500000$	10.601000	2.809400	-0.005800	-0.053867	
$t_3 = 4$	12.002800	2.803600	0.048567	-0.054367	0.001000
$t_4 = 2$	6.541300	2.730750			

TABLE 1.5. Différences divisées de f .

question 1 en rajoutant un cinquième point, en bas du tableau ; voir le tableau 1.5. On déduit de ce tableau la valeur de $f[t_0, t_1, t_2, t_3, t_4]$ donnée par :

$$f[t_0, t_1, t_2, t_3, t_4] = 0.001000.$$

On a alors

$$E_3(\tau) \approx 1.0 \cdot 10^{-3} \times (\tau - t_0) \dots (\tau - t_3) = 2.37599 \cdot 10^{-5}.$$

Exercices facultatifs

CORRECTION DE L'EXERCICE 1.6.

(1) On rappelle que Π_n vérifie

$$\forall i \in \{0, \dots, n\}, \quad \Pi_n(x_i) = y_i. \quad (1.54)$$

D'après la proposition A.2, le polynôme Π_n est de degré au plus n . Il suffit de choisir un polynôme P de degré p tel que

$$\forall i \in \{0, \dots, n\}, \quad y_i = P(x_i). \quad (1.55)$$

Le polynôme d'interpolation Π_n vaut alors P . Il est en effet de degré p , inférieur ou égal à n par définition et, d'après (1.55), (1.54) est vérifiée pour $\Pi_n = P$. P vérifie donc l'équation (A.5) de la proposition A.2. Par unicité du polynôme d'interpolation,

$$\Pi_n = P. \quad (1.56)$$

(2) Prenons par exemple

$$P(x) = x^2.$$

D'après (1.55), les y_i sont donnés par

$$\forall i \in \{0, \dots, n\}, \quad y_i = x_i^2 = i^2.$$

Il serait fastidieux de déterminer Π_3 . Il suffit d'utiliser (1.56) qui nous donne :

$$\Pi_3(x) = x^2.$$

CORRECTION DE L'EXERCICE 1.7.

On pourra aussi consulter la correction de l'exercice de TD 1.5 page 12, très proche de cet exercice.

(1)

x (en mètres)	0	100	1500	10000
$t(x)$ (en secondes)	0	13	245	1980

TABLE 1.6. Les données du sportif.

Parmi les différents points $(d_i)_{0 \leq i \leq 3}$ (les distances données dans la première ligne du tableau 1.6) on considère trois points successifs qui encadrent "au mieux" la donnée

$$d = 5000. \quad (1.57)$$

On utilise tout d'abord le fait que l'on fait de l'"inter"polation (la donnée d est à l'intérieur des distances $(d_i)_{0 \leq i \leq 3}$). De plus, l'écart entre d et les d_i doit être le plus faible possible. En outre, l'écart entre les différents d_i doit être le plus faible possible. Enfin, on admet que la situation doit être la plus symétrique possible. Finalement, on choisit donc :

$$X_0 = 100, \quad (1.58a)$$

$$X_1 = 1500, \quad (1.58b)$$

$$X_2 = 10000, \quad (1.58c)$$

et les durées correspondantes données par

$$Y_0 = 13, \quad (1.59a)$$

$$Y_1 = 245, \quad (1.59b)$$

$$Y_2 = 1980, \quad (1.59c)$$

Il était préférable d'utiliser la méthode de Newton, mais le calcul par les polynômes de Lagrange est aussi présenté.

(a) Chacun des polynômes de Lagrange l_i (de degré 2) est donné par la formule :

$$\forall i \in \{0, \dots, n\}, \quad l_i(X) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{X - X_j}{X_i - X_j}. \quad (1.60)$$

On a donc successivement

$$\begin{aligned} l_0(X) &= \frac{(X - 1500)(X - 10000)}{(100 - 1500)(100 - 10000)}, \\ l_1(X) &= \frac{(X - 100)(X - 10000)}{(1500 - 100)(1500 - 10000)}, \\ l_2(X) &= \frac{(X - 100)(X - 1500)}{(10000 - 100)(10000 - 1500)}. \end{aligned}$$

soit encore après calculs :

$$l_0(X) = 7.215007214 \cdot 10^{-8} X^2 - 8.297258297 \cdot 10^{-4} X + 1.082251081, \quad (1.61a)$$

$$l_1(X) = -8.403361345 \cdot 10^{-8} X^2 + 8.487394957 \cdot 10^{-4} X - 8.403361345 \cdot 10^{-2}, \quad (1.61b)$$

$$l_2(X) = 1.188354130 \cdot 10^{-8} X^2 - 1.901366606 \cdot 10^{-5} X + 1.782531193 \cdot 10^{-3}. \quad (1.61c)$$

Ensuite, le polynôme interpolateur de degré 2, Π_2 , est donné par la formule :

$$\Pi_2(X) = \sum_{i=0}^n Y_i l_i(X). \quad (1.62)$$

Ici, on a donc :

$$\Pi_2(X) = Y_0 l_0(X) + Y_1 l_1(X) + Y_2 l_2(X).$$

Après calculs, il vient :

$$\Pi_2(X) = 3.879127409 \cdot 10^{-6} X^2 + 1.595076819 \cdot 10^{-1} X - 2.989559460. \quad (1.63)$$

(b)

$X_i \setminus k$	0	1	2
$X_0 = 100$	13		
		$1.657142857 \cdot 10^{-1}$	
$X_1 = 1500$	245		$3.879127409 \cdot 10^{-6}$
		$2.041176470 \cdot 10^{-1}$	
$X_2 = 10000$	1980		

TABLE 1.7. Différences divisées de f .

Pour calculer le polynôme sous la forme de Newton, on détermine tout d'abord les différences divisées $f[X_i, \dots, X_{i+k}]$ données dans le tableau 1.7. Ensuite, on n'utilise plus que les différences divisées qui sont encadrées et le polynôme interpolateur est donné par la formule :

$$\Pi_2(X) = \sum_{i=0}^n f[X_0, \dots, X_i](X - X_0) \dots (X - X_{i-1}). \quad (1.64)$$

Ici, on a donc :

$$\Pi_2(X) = f[X_0] + f[X_0, X_1](X - X_0) + f[X_0, X_1, X_2](X - X_0)(X - X_1).$$

On a successivement

$$\begin{aligned} X - X_0 &= X - 10^2, \\ (X - X_0)(X - X_1) &= X^2 - 1.600000000 \cdot 10^3 X + 1.500000000 \cdot 10^5. \end{aligned}$$

Après calculs, on retrouve donc bien le polynôme déterminé par la méthode de Lagrange (voir équation (1.63)).

En prenant

$$d = 5000, \tag{1.65}$$

on en déduit la valeur de donnée par $\Pi_2(d)$ donnée par

$$\tau = \Pi_2(d) = 891.5270, \tag{1.66}$$

ce qui fournit une approximation de la performance sur le 5 000 m.

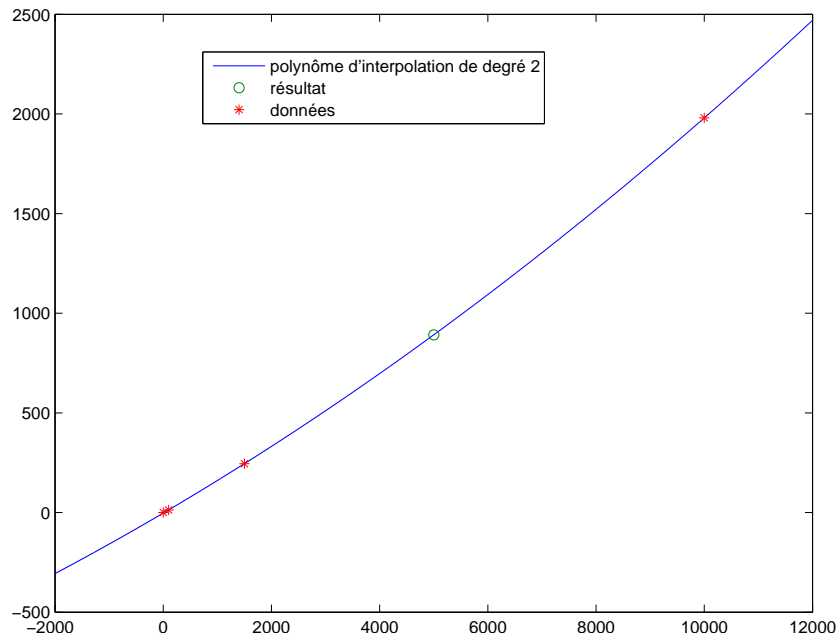


FIGURE 1.10. les données et le résultat.

Voir aussi la figure 1.10 où sont représentées les données, Π_2 , d et $\Pi_2(d)$.

REMARQUE 1.5. On peut aussi utiliser les trois premiers points, ce qui était un peu moins moins pertinent. D'une part, on ne fait plus de l'interpolation car le point est à l'extérieur des données. De plus, on utilise le temps nul correspondant à une distance nulle, ce qui n'est pas très intéressant.

Présentons néanmoins ce calcul. Il était préférable d'utiliser la méthode de Newton, mais le calcul par les polynômes de Lagrange est aussi présenté.

(a) Chacun des polynômes de Lagrange l_i (de degré 2) est donné par la formule :

$$\forall i \in \{0, \dots, n\}, \quad l_i(X) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{X - X_j}{X_i - X_j}. \tag{1.67}$$

On a donc successivement

$$l_0(X) = \frac{(X - 100)(X - 1500)}{(0 - 100)(0 - 1500)},$$

$$l_1(X) = \frac{(X)(X - 1500)}{(100)(100 - 1500)},$$

$$l_2(X) = \frac{(X)(X - 100)}{(1500)(1500 - 100)}.$$

soit encore après calculs :

$$l_0(X) = 6.666666667 \cdot 10^{-6} X^2 - 1.066666667 \cdot 10^{-2} X + 1, \quad (1.68a)$$

$$l_1(X) = -7.142857142 \cdot 10^{-6} X^2 + 1.071428571 \cdot 10^{-2} X, \quad (1.68b)$$

$$l_2(X) = 4.761904762 \cdot 10^{-7} X^2 - 4.761904762 \cdot 10^{-5} X. \quad (1.68c)$$

Ensuite, le polynôme interpolateur de degré 2, Π_2 , est donné par la formule :

$$\Pi_2(X) = \sum_{i=0}^n Y_i l_i(X). \quad (1.69)$$

Ici, on a donc :

$$\Pi_2(X) = Y_0 l_0(X) + Y_1 l_1(X) + Y_2 l_2(X).$$

Après calculs, il vient :

$$\Pi_2(X) = 2.380952381 \cdot 10^{-5} X^2 + 1.276190476 \cdot 10^{-1} X. \quad (1.70)$$

(b)

$X_i \setminus k$	0	1	2
$X_0 = 0$	0		
		$1.300000000 \cdot 10^{-1}$	
$X_1 = 100$	13		$2.380952381 \cdot 10^{-5}$
		$1.657142857 \cdot 10^{-1}$	
$X_2 = 1500$	245		

TABLE 1.8. Différences divisées de f .

Pour calculer le polynôme sous la forme de Newton, on détermine tout d'abord les différences divisées $f[X_i, \dots, X_{i+k}]$ données dans le tableau 1.8. Ensuite, on n'utilise plus que les différences divisées qui sont encadrées et le polynôme interpolateur est donné par la formule :

$$\Pi_2(X) = \sum_{i=0}^n f[X_0, \dots, X_i](X - X_0) \dots (X - X_{i-1}). \quad (1.71)$$

Ici, on a donc :

$$\Pi_2(X) = f[X_0] + f[X_0, X_1](X - X_0) + f[X_0, X_1, X_2](X - X_0)(X - X_1).$$

On a successivement

$$X - X_0 = X,$$

$$(X - X_0)(X - X_1) = X^2 - 10^2 X.$$

Après calculs, on retrouve donc bien le polynôme déterminé par la méthode de Lagrange (voir équation (1.70)).

On en déduit la valeur de donnée par $\Pi_2(d)$ donnée par

$$\tau = \Pi_2(d) = 1233.3333, \quad (1.72)$$

ce qui fournit une approximation de la performance sur le 5 000 m à comparer avec la valeur donnée par (1.66).

(2) L'expression analytique de l'erreur est donnée par

$$E_2(x) = \frac{f^{(3)}(\xi)}{3!} \omega_3(x), \quad (1.73)$$

où ici f et ξ ne sont pas connus. Cette formule provient de la propriété A.4 utilisée ici avec $n = 2$.

Il est en théorie impossible de calculer une telle erreur. Cependant, on peut supposer par exemple que f est polynomiale de degré 3, ou ce qui revient au même que $f^{(3)}$ ne varie pas. On a alors une approximation de l'erreur, obtenue en prenant un point supplémentaire :

$$E_2(x) \approx f[X_0, \dots, X_{n+1}](x - X_0) \dots (x - X_n), \quad (1.74)$$

utilisée ici avec $n = 2$. Cela provient de (1.73); en effet, si f est polynomiale de degré 3, le coefficient dominant de f est $f[t_0, \dots, t_{n+1}]$ puisque c'est le coefficient dominant de $f = \Pi_3$ sur la forme de Newton. On a alors pour tout ξ , puisque $f^{(3)}$ est constant :

$$f^{(3)}(\xi) = 3!f[X_0, \dots, X_3],$$

ce qui donne, réinjecté dans (1.73) :

$$E_2(x) = f[X_0, \dots, X_3] \omega_3(x).$$

Si cette fois-ci, $f^{(3)}$ varie peu, cela devient une approximation et on a donc, dans tous les cas (1.73). Prenons $X_0 = 100$, $X_1 = 1500$, $X_2 = 10000$ et $X_3 = 0$. On complète le tableau déjà fait dans la

$X_i \setminus k$	0	1	2	3
$X_0 = 100$	13			
$X_1 = 1500$	245	$1.657142857 \cdot 10^{-1}$	$3.879127409 \cdot 10^{-6}$	
$X_2 = 10000$	1980	$2.041176470 \cdot 10^{-1}$	$4.078431372 \cdot 10^{-6}$	$-1.993039639 \cdot 10^{-9}$
$X_3 = 0$	0	$1.980000000 \cdot 10^{-1}$		

TABLE 1.9. Différences divisées de f .

question 1 en rajoutant un quatrième point; voir le tableau 1.9. On déduit de ce tableau la valeur de $f[X_0, X_1, X_2, X_3]$ donnée par :

$$f[X_0, X_1, X_2, X_3] = -1.993039639 \cdot 10^{-9}.$$

On a alors

$$E_2(d) \approx -1.993039639 \cdot 10^{-9} \times (d - X_0)(d - X_2)(d - X_2) = 170.903149138443,$$

ce qui est assez élevé dans l'absolue ; l'erreur relative est donnée par exemple par (en utilisant la valeur de τ donnée par (1.66))

$$e_2(d) = \frac{E_2(d)}{\tau} \approx 0.19169710 = 19.169710\%,$$

ce qui est "plus raisonnable".

REMARQUE 1.6. On peut tenir le raisonnement suivant Pour calculer l'erreur

$$E_2(x) = f(x) - \Pi_2(x) = \Pi_3(x) - \Pi_2(x),$$

on détermine $\Pi_3(x)$ grâce au tableau 1.9 page ci-contre. On écrit alors

$$E_2(d) = \Pi_3(d) - \Pi_2(d),$$

en utilisant la valeur de $\Pi_2(d)$ donnée par (1.66). Mais, c'est beaucoup plus long et identique ! On sait en effet d'après les relations de récurrence de Π_n que :

$$\Pi_3(d) = \Pi_2(d) + f[X_0, X_1, X_2, X_3](d - X_0)(d - X_1)(d - X_2),$$

et donc on a

$$E_2(d) = \Pi_3(d) - \Pi_2(d) = f[X_0, X_1, X_2, X_3](d - X_0)(d - X_1)(d - X_2),$$

et c'est donc exactement identique à (1.74) avec $n = 2$ au point d !

REMARQUE 1.7. Il serait peut-être pertinent de remettre en cause ce calcul, non pas sur le plan mathématique, mais sur le plan sportif ou biomécanique : est-il pertinent de se servir des performances sur des durées différentes et donc des compétences différentes ?

Intégration

CORRECTION DE L'EXERCICE 2.1.

Pour tout cet exercice, découpons $[A, B]$ en sous-intervalles à pas constant h ($h \in \mathbb{R}_+^*$), notés $[x_i, x_{i+1}]$. Ainsi

$$x_0 = A; \quad x_N = B; \quad \forall i \in \{0, \dots, N-1\} \quad x_{i+1} - x_i = h \quad \text{d'où} \quad h = \frac{B-A}{N}. \quad (2.1)$$

Par suite pour tout i de $\{0, \dots, N\}$

$$x_i = A + ih. \quad (2.2)$$

(1) L'approximation de l'intégrale par la méthode des trapèzes composite (voir tableau A.4) est

$$I_N^T = \frac{h}{2}(f(A) + f(B)) + h \sum_{i=1}^{N-1} f(x_i).$$

Pour $N = 4$, on obtient

$$I_N^T = 1/8 + 1/8 e^{-1} + 1/4 e^{-1/16} + 1/4 e^{-1/4} + 1/4 e^{-9/16}.$$

et donc

$$I_N^T = 0.742984. \quad (2.3)$$

(2) L'erreur commise par la méthode des trapèzes composite (voir tableau A.6) est

$$E_N^T = -h^2 \frac{(B-A)}{12} f''(\eta) \quad \text{avec } \eta \in [A, B].$$

En théorie (et si $f''(\eta)$ ne varie que peu), si h est divisé par 2, cette erreur est divisée par 4.

REMARQUE 2.1. On peut aussi écrire

$$|E_N^T| \leq h^2 \frac{(B-A)}{12} \max_{x \in [A, B]} |f''(x)|,$$

et remarquer que cette dernière quantité est divisée par 4 quand h est divisée par 2.(3) Notons $M_2 = \sup_{x \in [A, B]} |f''(x)|$. On a

$$f''(x) = 2e^{-x^2}(-1 + 2x^2).$$

que l'on peut majorer par $M_2 = 6$, comme on a déjà fait l'exercice 1.4 page 10 (voir (1.37)). Pour avoir une erreur inférieure à $\varepsilon > 0$, il suffit donc que

$$h^2 \frac{(B-A)}{12} M_2 \leq \varepsilon, \quad (2.4)$$

ce qui est équivalent à

$$\frac{(B-A)^3}{12N^2} M_2 \leq \varepsilon,$$

ou encore

$$\sqrt{\frac{(B-A)^3}{12\varepsilon}} M_2 \leq N. \quad (2.5)$$

Numériquement, on obtient donc pour $A = 0$, $B = 1$, et $\varepsilon = 10^{-4}$

$$N \geq 70.710678$$

et puisque N est entier, on choisit

$$N = 71 \tag{2.6}$$

REMARQUE 2.2. On peut être plus subtil et utiliser la majoration de la remarque 1.3 page 11 de l'exercice 1.4 (voir (1.44)) et on obtient $N = 41$, ce qui est plus faible que (2.6).

En utilisant le calcul symbolique de matlab, on peut avoir une approximation très précise de l'intégrale recherchée :

$$I = 0.746824, \tag{2.7}$$

Pour la valeur de N donnée par (2.6), on a

$$I_N^T = 0.746812, \tag{2.8}$$

on vérifie *a posteriori* alors que

$$|I - I_N^T| = 1.2162 \cdot 10^{-5}.$$

ce qui est bien inférieure à 10^{-4} .

- (4) L'approximation de l'intégrale par la méthode de Simpson composite (voir tableau A.4) est

$$I_N^S = \frac{h}{6} \left(f(A) + f(B) + 2 \sum_{i=1}^{N-1} f(x_i) + 4 \sum_{i=0}^{N-1} f\left(x_i + \frac{h}{2}\right) \right).$$

Pour $N = 2$, on obtient

$$I_N^T = 1/12 + 1/12 e^{-1} + 1/6 e^{-1/4} + 1/3 e^{-1/16} + 1/3 e^{-9/16}.$$

et donc

$$I_N^T = 0.746855. \tag{2.9}$$

- (5) L'erreur commise par la méthode de Simpson composite (voir tableau A.6) est

$$E_N^S = -h^4 \frac{(B-A)}{2880} f^{(4)}(\eta) \text{ avec } \eta \in [A, B].$$

Comme précédemment, on pose $M_4 = \sup_{x \in [A, B]} |f^{(4)}(x)|$. On a

$$f^{(4)}(x) = 4 e^{-x^2} (3 - 12x^2 + 4x^4). \tag{2.10}$$

Pour majorer $|f^{(4)}(x)|$, on procède exactement comme dans l'exercice 1.4 page 10. On écrit

$$f^{(4)}(x) = 4e^{-x^2} p_4(x), \tag{2.11}$$

où

$$p_4(x) = 3 - 12x^2 + 4x^4. \tag{2.12}$$

Puis on raisonne comme dans l'inégalité (1.31) :

$$\forall x \in [0, 1], \quad |f^{(4)}(x)| \leq 4 \left| e^{-x^2} \right| |p_4(x)|, \tag{2.13}$$

et, comme dans les inégalités (1.32) et (1.33), on a

$$\forall x \in [0, 1], \quad |f^{(4)}(x)| \leq 4e^{-x^2} |p_4(x)|, \tag{2.14}$$

et

$$\forall x \in [0, 1], \quad |f^{(4)}(x)| \leq 4 |p_4(x)|. \tag{2.15}$$

Pour majorer $|p_4(x)|$, on raisonne comme on a déjà fait dans les inégalités (1.34), (1.35) et (1.36) en utilisant l'inégalité triangulaire (1.30), écrite ici sous sa forme plus générale :

$$\forall (b_1, \dots, b_q) \in \mathbb{R}^q, \quad \left| \sum_{i=1}^q b_i \right| \leq \sum_{i=1}^q |b_i|. \quad (2.16)$$

On écrit, ensuite, pour tout polynôme Q :

$$Q(x) = \sum_{i=0}^p a_i x^i,$$

puis les inégalités successives (grâce à (2.16))

$$\begin{aligned} |Q(x)| &= \left| \sum_{i=0}^p a_i x^i \right|, \\ &\leq \sum_{i=0}^p |a_i x^i|, \\ &\leq \sum_{i=0}^p |a_i| |x^i|, \\ &\leq \sum_{i=0}^p |a_i| |x|^i, \end{aligned}$$

et puisque $|x| \leq 1$

$$\leq \sum_{i=0}^p |a_i|.$$

On a donc

$$\forall x \in [0, 1], \quad |Q(x)| \leq \sum_{i=0}^p |a_i|.$$

Ici, pour p_4 donné par (2.12), dont les coefficients sont $\{4, 0, -12, 0, 3\}$, on a

$$\forall x \in [0, 1], \quad |p_4(x)| \leq 19.$$

Ainsi, grâce à (2.15), on a

$$\forall x \in [0, 1], \quad |f^{(4)}(x)| \leq M_4. \quad (2.17)$$

où

$$M_4 = 76. \quad (2.18)$$

Pour avoir une erreur inférieure à $\varepsilon > 0$, il suffit donc que

$$h^4 \frac{(B-A)}{2880} M_4 \leq \varepsilon. \quad (2.19)$$

On raisonne comme dans la question 3. Il faut avoir

$$N \geq \sqrt[4]{\frac{(B-A)^5}{2880\varepsilon} M_4}. \quad (2.20)$$

Numériquement, on obtient pour $A = 0$, $B = 1$, et $\varepsilon = 10^{-4}$

$$N \geq 4.030466,$$

et puisque N est entier, on choisit

$$N = 5, \quad (2.21)$$

à comparer avec la valeur donnée par (2.6).

REMARQUE 2.3. Si utilise la fonction `maxabsfun`, fournie sur le site habituel, pour déterminer avec plus de soin, le majorant exact (puisqu'atteint en un point) de g , définie par (2.10), on obtient

$$M_4 = 12,$$

qui est un peu plus précis que (2.18). On pourra consulter l'annexe C pour une preuve rigoureuse de ce résultat. On peut donc remplacer (2.21) par

$$N = 3.$$

CORRECTION DE L'EXERCICE 2.2.

- (1) Par linéarité des fonctions $f \mapsto I(f)$ et $f \mapsto Q(f)$, l'exactitude de la formule de quadrature sur l'espace vectoriel des fonctions polynomiales de degré inférieur à 2 est équivalente à l'exactitude de la formule de quadrature pour les vecteurs de la base canonique de cet espace vectoriel, c'est-à-dire pour les fonctions définies par $f(x) = 1$, $f(x) = x$ et $f(x) = x^2$. Cela donne le système linéaire suivant

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1/2 \\ 1/3 \end{pmatrix}.$$

On peut le résoudre à la main ou matriciellement pour obtenir :

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{bmatrix} 2/3 \\ 1/3 \\ 1/6 \end{bmatrix}$$

La formule est au moins de degrés d'exactitude 2. Pour vérifier si elle est degré d'exactitude exactement 2, on calcule l'erreur correspondant à un polynôme de degré 3. Pour cela par linéarité, on calcule $I(x^3) - Q(x^3)$. Si cette erreur est nulle, le degré est au moins 3. Sinon, il est exactement 2, ce qui est le cas ici.

- (2)

$$I \approx 0.789293,$$

que l'on peut comparer à (2.3), (2.7), (2.8) et (2.9).

Exercices facultatifs

Cet exercice a été donné à l'examen de MNB en Mécanique à l'automne 2017.

CORRECTION DE L'EXERCICE 2.3.

On renverra aussi à l'exercice de TD 2.2 proche de cet exercice !

Pour toute la suite, les nœuds interpolant sont notés x_0, \dots, x_n .

- (1) (a) Par linéarité des fonctions $f \mapsto I(f)$ et $f \mapsto Q(f)$, l'exactitude de la formule de quadrature sur l'espace vectoriel des fonctions polynomiales de degré inférieur à 2, de dimension 3, est équivalente à l'exactitude de la formule de quadrature pour les vecteurs de la base canonique de cet espace vectoriel, c'est-à-dire pour les fonctions définies par

$$e_0(x) = 1, \quad e_1(x) = x, \quad e_2(x) = x^2.$$

Cela donne le système linéaire suivant

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1/2 & 1 \\ 0 & 1/4 & 1 \end{pmatrix} C = \begin{pmatrix} 1 \\ 1/2 \\ 1/3 \end{pmatrix}, \quad (2.22)$$

où C est le vecteur des coefficients W_i recherchés. On peut le résoudre à la main ou matriciellement pour obtenir :

$$C = \begin{pmatrix} 1/6 \\ 2/3 \\ 1/6 \end{pmatrix}. \quad (2.23)$$

REMARQUE 2.4.

La matrice intervenant dans le système linéaire (2.22) est en fait la matrice de Vandermonde suivante, correspondant aux points x_i donnés par

$$h = (b - a)/n, \quad \forall i \in \{0, \dots, n\}, \quad x_i = a + ih, \quad (2.24)$$

avec, ici $n = 2$ et $a = 0, b = 1$:

$$D_n = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ x_0 & x_1 & x_2 & \dots & x_{n-1} & x_n \\ x_0^2 & x_1^2 & x_2^2 & \dots & x_{n-1}^2 & x_n^2 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ x_0^{n-1} & x_1^{n-1} & x_2^{n-1} & \dots & x_{n-1}^{n-1} & x_n^{n-1} \\ x_0^n & x_1^n & x_2^n & \dots & x_{n-1}^n & x_n^n \end{pmatrix}. \quad (2.25)$$

Notons que cette matrice peut s'inverser en théorie en utilisant les polynômes de Lagrange! Voir [BM03, Exercice 2.5 p. 55]. Notons que dans le cas $n = 2$, l'inverse explicite a été calculé dans cet exercice :

$$D_2^{-1} = \begin{pmatrix} \frac{x_1 x_2}{(x_0 - x_1)(x_0 - x_2)} & -\frac{x_1 + x_2}{(x_0 - x_1)(x_0 - x_2)} & \frac{1}{(x_0 - x_1)(x_0 - x_2)} \\ \frac{x_0 x_2}{(x_1 - x_0)(x_1 - x_2)} & -\frac{x_0 + x_2}{(x_1 - x_0)(x_1 - x_2)} & \frac{1}{(x_1 - x_0)(x_1 - x_2)} \\ \frac{x_0 x_1}{(x_2 - x_0)(x_2 - x_1)} & -\frac{x_0 + x_1}{(x_2 - x_0)(x_2 - x_1)} & \frac{1}{(x_2 - x_0)(x_2 - x_1)} \end{pmatrix},$$

où les x_i sont donnés par (2.24). Ici, on a explicitement

$$D_2^{-1} = \begin{pmatrix} 1 & -3 & 2 \\ 0 & 4 & -4 \\ 0 & -1 & 2 \end{pmatrix}. \quad (2.26)$$

Si on veut l'inverser directement, on notera que cette matrice est sensible à l'inversion numérique (car elle a un conditionnement important quand n grandit). Il est préférable d'utiliser la méthode de la question 1b pour la calculer pour n grand.

- (b) (i) Pour calculer $\Pi_2(f)$, le polynôme d'interpolation de f aux nœuds $0, 1/2, 1$ en fonction de $f(0), f(1/2), f(1)$, on utilise la forme de Newton, en prenant bien soin de laisser les valeurs de $f(0), f(1/2), f(1)$ générique, comme le montre la suite.

Pour calculer le polynôme sous la forme de Newton, on détermine tout d'abord les différences divisées $f[x_i, \dots, x_{i+k}]$ données dans le tableau 2.1. Ensuite, on n'utilise plus que les différences divisées qui sont encadrées et le polynôme interpolateur de degré 2, $\Pi_2(f)$, est donné par la formule :

$$\Pi_2(f)(x) = \sum_{i=0}^n f[x_0, \dots, x_i](x - x_0) \dots (x - x_{i-1}). \quad (2.27)$$

$x_i \setminus k$	0	1	2
$x_0 = 0$	$f(0)$		
$x_1 = 1/2$	$f(1/2)$	$2f(1/2) - 2f(0)$	$2f(1) - 4f(1/2) + 2f(0)$
$x_2 = 1$	$f(1)$	$2f(1) - 2f(1/2)$	

TABLE 2.1. Différences divisées de f .

Ici, on a donc :

$$\Pi_2(f)(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1).$$

On a successivement

$$\begin{aligned} x - x_0 &= x, \\ (x - x_0)(x - x_1) &= x^2 - 1/2x. \end{aligned}$$

Après calculs, il vient :

$$\Pi_2(f)(x) = 2x^2f(1) - 4x^2f(1/2) + 2x^2f(0) + 4xf(1/2) - 3xf(0) - xf(1) + f(0). \quad (2.28)$$

(ii) L'intégrale $\mathcal{I}(\Pi_2(f))$ définie par

$$\mathcal{I}(\Pi_2(f)) = \int_0^1 \Pi_2(f)(x)dx, \quad (2.29)$$

s'obtient en intégrant le polynôme $\Pi_2(f)$ qui vient d'être déterminé. En intégrant les fonctions 1, x et x^2 , sur $[0, 1]$, on obtient donc finalement

$$\mathcal{I}(\Pi_2(f)) = 1/6f(1) + 2/3f(1/2) + 1/6f(0). \quad (2.30)$$

(iii) On cherche à trouver les coefficients W_i de telle sorte que la formule de quadrature intègre exactement les polynômes jusqu'au degré 2. Cela est donc équivalent à ce qu'elle soit exacte pour $\Pi_2(f)$, pour toute fonction f , puisque $\Pi_2(f)$ est un polynôme de degré au plus 2. C'est donc équivalent à l'égalité de $Q(\Pi_2(f))$ et de $\mathcal{I}(\Pi_2(f))$. Dans cette dernière égalité, chacune des valeurs de $\Pi_2(f)(x_i)$ est remplacée par définition par $f(x_i)$. D'après (2.30), on voit donc apparaître les coefficients W_i qui correspondent bien à ceux donnés par (2.23).

(2) Pour déterminer le degré d'exactitude (que l'on appelle aussi l'ordre) de la méthode, nous avons deux méthodes.

(a) Soit, on remarque que, en posant $a = 0$ et $b = 1$, la formule de quadrature établie (2.30) est équivalente à

$$Q(f) = \frac{1}{6}(b-a)(f(a) + 4f((a+b)/2) + f(b)),$$

et on reconnaît la méthode de Simpson, dont on connaît l'erreur, d'après les formules rappelées dans l'énoncé. Voir tableau 2.2. Ainsi, ici l'erreur est proportionnelle à $f^{(4)}(\eta)$, et donc nulle pour les polynômes de degrés au plus trois, et non nulle, par exemple, pour le polynôme x^4 . Ainsi, le degré d'exactitude (appelé aussi l'ordre) de la méthode, qui correspond au plus haut degré du polynôme exactement intégré par la formule de quadrature, vaut donc 3.

méthode	erreur
rectangle	$\frac{(b-a)^2}{2} f'(\eta)$
milieu	$\frac{(b-a)^3}{24} f''(\eta)$
trapèze	$-\frac{(b-a)^3}{12} f''(\eta)$
Simpson	$-\frac{(b-a)^5}{2880} f^{(4)}(\eta)$

TABLE 2.2. Erreurs pour les méthodes élémentaires sur $[a, b]$; η appartient à $]a, b[$.

- (b) Si on ne reconnaît pas la méthode vue en cours, il suffit de vérifier que le plus haut degré du polynôme exactement intégré par la formule de quadrature vaut 3. Ainsi, le degré d'exactitude (que l'on appelle aussi l'ordre) de la méthode est 3.
- (3) (a) Pour donner une approximation numérique de l'intégrale, on utilise la formule de quadrature déterminée; on a donc

$$\begin{aligned} \mathcal{I} &= \begin{pmatrix} 1/6 \\ 2/3 \\ 1/6 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ e^{-1/4} \\ e^{-1} \end{pmatrix}, \\ &= 1/6 + 2/3 e^{-1/4} + 1/6 e^{-1}, \\ &\approx 0.747180428909510. \end{aligned}$$

- (b) On a ici $a = 0$ et $b = 1$. D'après le tableau donné dans l'énoncé (voir tableau 2.2), l'erreur de la méthode de quadrature est donnée par

$$E(f) = \beta f^{(4)}(\xi), \text{ où } \xi \in [0, 1], \quad (2.31)$$

avec, ici,

$$\beta = -\frac{1}{2880}. \quad (2.32)$$

On détermine alors la dérivée à l'ordre exigé :

$$f^{(4)}(x) = 4 e^{-x^2} (3 - 12 x^2 + 4 x^4).$$

On majore la valeur absolue de cette fonction sur $[0, 1]$ par M donné par

$$M = 12 \approx 12. \quad (2.33)$$

Ici, cette valeur a été obtenue grâce à la fonction fournie `maxabsfun`. Compte tenu de (2.31), (2.32) et (2.33), on a donc une erreur finalement majorée par

$$\varepsilon = \frac{1}{240} \approx 4.166667 \cdot 10^{-3}. \quad (2.34)$$

On peut aussi déterminer la valeur exacte (par une approximation très précise!) : on obtient

$$\mathcal{I} = 0.746824132812427,$$

et donc une erreur réellement commise de

$$E = 3.562961 \cdot 10^{-4},$$

qui est bien inférieure à celle donnée par (2.34).

REMARQUE 2.5.

La façon de déterminer les coefficients de la forme de quadrature de la question 1b est la plus efficace. C'est cette méthode qui permet à la fois de déterminer les méthodes élémentaires usuelles (rectangles, trapèzes, ...) mais aussi l'erreur de convergence associée. Voir par exemple [BM03, Chapitre 3, p. à 81 à 89].

REMARQUE 2.6.

En raisonnant comme dans la remarque 2.4, une autre façon de déterminer les poids W_i est de remarquer que, pour un polynôme p de degré au plus $n = 2$,

$$\int_0^1 p(x)dx = \int_0^1 \sum_{i=0}^2 p(x_i)l_i(x)dx, \\ = \sum_{i=0}^2 p(x_i) \int_0^1 l_i(x)dx,$$

où les l_i sont les polynômes de Lagrange, et donc

$$\forall i \in \{0, \dots, n\}, \quad W_i = \int_0^1 l_i(x)dx.$$

Par exemple, on a

$$W_0 = \int_0^1 2x^2 - 3x + 1dx, \\ = 1/6,$$

ce qui est le premier coefficient déjà déterminé.

REMARQUE 2.7.

valeurs de n	nom	coefficients	ordre
0	rectangle	1	0
1	trapèze	1/2, 1/2	1
2	Simpson	1/6, 2/3, 1/6	3
3	Simpson 3/8	1/8, 3/8, 3/8, 1/8	3
4	Boole-Villarceau	$\frac{7}{90}, \frac{16}{45}, 2/15, \frac{16}{45}, \frac{7}{90}$	5
5		$\frac{19}{288}, \frac{25}{96}, \frac{25}{144}, \frac{25}{144}, \frac{25}{96}, \frac{19}{288}$	5
6	Weddle-Hardy	$\frac{41}{840}, \frac{9}{35}, \frac{9}{280}, \frac{34}{105}, \frac{9}{280}, \frac{9}{35}, \frac{41}{840}$	7
7		$\frac{751}{17280}, \frac{3577}{17280}, \frac{49}{640}, \frac{2989}{17280}, \frac{49}{640}, \frac{3577}{17280}, \frac{751}{17280}$	7
8		$\frac{989}{28350}, \frac{2944}{14175}, -\frac{464}{14175}, \frac{5248}{14175}, -\frac{454}{2835}, \frac{5248}{14175}, -\frac{464}{14175}, \frac{2944}{14175}, \frac{989}{28350}$	9
9		$\frac{2857}{89600}, \frac{15741}{89600}, \frac{27}{2240}, \frac{1209}{5600}, \frac{2889}{44800}, \frac{2889}{44800}, \frac{1209}{5600}, \frac{27}{2240}, \frac{15741}{89600}, \frac{2857}{89600}$	9

TABLE 2.3. Noms, coefficients et ordres des 10 premières méthodes (sur l'intervalle $[0, 1]$).

De façon plus générale, la méthode de quadrature étudiée, pour n quelconque est appelée la méthode de Newton-Cotes fermée. On peut calculer les coefficients sur tout intervalle $[a, b]$. Les coefficients et les ordres pour les premières valeurs de n sont donnés dans le tableau 2.3. Pour obtenir les coefficients sur un autre intervalle $[a, b]$ quelconque, il suffit de multiplier chacun de ces coefficients par $b - a$. Les points x_i correspondant sont naturellement donnés par $x_0 = 0$ si $n = 0$ et sinon

$$\forall i \in \{0, \dots, n\}, \quad x_i = a + ih \text{ où } h = (b - a)/n.$$

En pratique, on choisit des valeurs de n pas trop élevées et souvent la méthode de Simpson suffira largement. De plus, au-delà, d'une certaine valeur de n_0 de n (pour $n_0 = 8$) les coefficients changent de signe ce qui favorise les propagations d'erreur d'arrondis.

REMARQUE 2.8.

On peut montrer que toute méthode de Newton-Cotes fermée à $n + 1$ (avec n non nul) points de support est d'ordre n si n est impair et d'ordre $n + 1$ si n est pair. Cela généralise les propriétés vues en cours (voir par exemple [BM03, remarque 3.20]) : la méthode du trapèze ($n = 1$) est d'ordre 1, celle de Simpson ($n = 2$) est d'ordre 3. On peut aussi montrer que comme pour les méthodes déjà vues en cours (comme dans le tableau 2.2) une méthode de Newton-Cotes d'ordre m et à $n + 1$ points possède une erreur $E_n(f)$ qui vérifie : il existe $\alpha_n > 0$ tel que, pour tout a, b , pour toute fonction $f \in C^{m+1}([a, b])$,

$$E_n(f) = \alpha_n (b - a)^{m+2} f^{(m+1)}(\xi), \text{ où } \xi \in [a, b]. \quad (2.35)$$

Plus précisément, $m = n$ si n est impair et $m = n + 1$ si n est pair, c'est-à-dire : si n est pair (non nul) il existe $\alpha_n > 0$ tel que, pour tout a, b , pour toute fonction $f \in C^{n+2}([a, b])$,

$$E_n(f) = \alpha_n (b - a)^{n+3} f^{(n+2)}(\xi), \text{ où } \xi \in [a, b], \quad (2.36)$$

et si n est impair, il existe $\alpha_n > 0$ tel que, pour tout a, b , pour toute fonction $f \in C^{n+1}([a, b])$,

$$E_n(f) = \alpha_n (b - a)^{n+2} f^{(n+1)}(\xi), \text{ où } \xi \in [a, b]. \quad (2.37)$$

Par exemple, d'après le tableau 2.2, pour $n = 1$, on a $\alpha_1 = -1/12$ et pour $n = 2$, on a $\alpha_2 = -1/2880$. De façon plus générale, on peut calculer les constantes α_n en remarquant que l'erreur commise pour le polynôme x^{m+1} (si m est l'ordre de la formule et n le degré non nul) vaut

$$\tilde{E}_n = \int_a^b x^{m+1} dx - \sum_{i=0}^n W_i x_i^{m+1}, \quad (2.38)$$

cette erreur non nulle pouvant être calculée informatiquement. En utilisant (2.35), cette erreur vaut aussi

$$\tilde{E}_n = \alpha_n (b - a)^{m+2} (m + 1)!. \quad (2.39)$$

Bref, grâce à (2.39), on a l'expression explicite et algébrique de α_n , pour n non nul :

$$\alpha_n = \frac{\tilde{E}_n}{(b - a)^{m+2} (m + 1)!}, \quad (2.40)$$

où \tilde{E}_n est déterminée de façon informatique grâce à (2.38). Si $n = 0$, on obtient α_0 grâce au tableau 2.2.

Voir par exemple le tableau 2.4.

REMARQUE 2.9.

Pour plus de détails sur les formules de Newton-Cotes, on pourra consulter [CM84] et plus particulièrement le chapitre 2 et l'exercice 2.3 ou encore

http://fr.wikipedia.org/wiki/Formule_de_Newton-Cotes

http://utbmjb.chez-alice.fr/UTBM/mt40/medianMT40_A03.zip

http://utbmjb.chez-alice.fr/UTBM/mt40/mediancorrigeMT40_A03.zip

valeurs de n	α_n
0	$1/2$
1	$-1/12$
2	$-\frac{1}{2880}$
3	$-\frac{1}{6480}$
4	$-\frac{1}{1935360}$
5	$-\frac{1}{37800000}$
6	$-\frac{1}{1567641600}$
7	$-\frac{167}{426924691200}$
8	$-\frac{37}{62783697715200}$
9	$-\frac{173}{458209960750080}$

TABLE 2.4. Coefficients α_n des 10 premières méthodes.

Équations non linéaires

CORRECTION DE L'EXERCICE 3.1.

- (1) La dérivée de la fonction
- f
- vaut

$$f'(x) = 2e^{2x} - 2,$$

strictement positive sur \mathbb{R}_+^* . Ainsi f est strictement croissante sur \mathbb{R}_+ . Puisque $f(0) = -7 < 0$ et que $\lim_{x \rightarrow +\infty} f(x) = \infty$, la fonction f est une bijection de \mathbb{R}_+ sur $[-7, \infty]$ et n'a donc qu'une seule racine sur \mathbb{R}_+ .

- (2) (a) On pose $a = 0$ et $b = 1.2$. Puisque $f(a)f(b) = -4.4 < 0$ et que f est continue, la méthode de la bisection est convergente vers une racine de f qui est dans \mathbb{R}_+ . D'après ce qui précède, cette racine est unique. Ainsi, sur cet intervalle, la méthode de la bisection est convergente vers l'unique racine de f sur \mathbb{R}_+ .

(b)

n	x_n	a_n	b_n
0	0.6000	0.0000	1.2000
1	0.9000	0.6000	1.2000
2	1.0500	0.9000	1.2000
3	1.1250	1.0500	1.2000
4	1.1625	1.1250	1.2000

TABLE 3.1. Valeurs des extrémités a_n et b_n des intervalles et des milieux x_n

Les résultats sont donnés dans le tableau 3.1.

- (3) On pose $a = 1.0$ et $b = 2.0$. Puisque $f(a)f(b) = -111.22138 < 0$, la méthode de la bisection est convergente vers l'unique racine de f . D'après la formule (A.26), on a $n = 27$.
- (4) (a) On a $f'(x) = 1 + \sin(x) \geq 0$, puisque $\sin(x) \geq -1$. Ainsi, la dérivée de f est positive sur \mathbb{R} et ne s'annule qu'aux points isolés $x_n = -\pi/2 + 2n\pi$ où n décrit \mathbb{Z} . Ainsi, la fonction f est strictement croissante sur \mathbb{R} et *a fortiori* sur $[-1.0, 2.0]$. Ainsi, la fonction f est une bijection de $[-1.0, 2.0]$ sur $[-1.540302, 2.416147]$ et n'a donc qu'une seule racine sur $[-1.0, 2.0]$.
- (b) (i) Puisque $f(a)f(b) = -3.72160 < 0$, la méthode de la bisection est convergente vers l'unique racine de f .
- (ii)

Les résultats sont donnés dans le tableau 3.2.

- (c) D'après la formule (A.26), on a
- $n = 29$
- .

CORRECTION DE L'EXERCICE 3.2.

n	x_n	a_n	b_n
0	0.5000	-1.0000	2.0000
1	1.2500	0.5000	2.0000
2	0.8750	0.5000	1.2500
3	0.6875	0.5000	0.8750
4	0.7813	0.6875	0.8750

TABLE 3.2. Valeurs des extrémités a_n et b_n des intervalles et des milieux x_n

- (1) Notons $f(x) = x^4 + 6x^2 - 60x + 36$. Le graphique semble nous montrer que la fonction est strictement décroissante puis strictement croissante, ce qu'on peut montrer facilement¹. Pour montrer que cette équation n'a que deux racines, il suffit alors d'exhiber une valeur pour laquelle la fonction s'annule. Par exemple pour $t_0 = 2$, $f(t_0) = -44 < 0$.
- (2) (a) Voir [BM03, Correction de l'exercice 4.7 p. 305].
 (b) Voir [BM03, Correction de l'exercice 4.7 p. 305].

Précisons la divergence de la méthode du point fixe. Pour ce cas-là, on peut utiliser la proposition A.17.

- On constate que le dénominateur D de g_2 , donné par

$$D(x) = x^3 + 6x - 60,$$

s'annule pour une seule valeur de x donnée² par

$$\beta = \sqrt[3]{30 + 2\sqrt{227}} - 2 \frac{1}{\sqrt[3]{30 + 2\sqrt{227}}} = 3.4072630664966. \quad (3.1)$$

Aussi, g_2 possède une limite infinie en β . Par ailleurs, D a une dérivée qui est un polynôme du second degré, sans racine réelle. Il est donc de signe constant, de celui de 1. Ainsi D est strictement croissant sur \mathbb{R} . Puisqu'il ne s'annule qu'en β et qu'il est de coefficient dominant strictement positif, D est donc strictement croissant et strictement positif sur $[\beta, +\infty[$ et croissant et strictement négatif sur $] - \infty, \beta]$. Ainsi, $36/D$ est strictement décroissant et strictement positif sur $[\beta, +\infty[$ et strictement décroissant et strictement négatif sur $] - \infty, \beta]$ et donc g_2 est strictement croissante et strictement négative sur $[\beta, +\infty[$ et strictement croissante et strictement positive sur $] - \infty, \beta]$. Voir le tableau de variation 3.3 page ci-contre de g_2 . Cela est confirmé par le graphe de g_2 , sur la figure 1(a).

- Sur l'intervalle $[3, \frac{31}{10}]$, on constate que g'_2 semble être comprise entre 5.2800000000 et 9.3039282563, ce qui est confirmé par le graphe 1(b). On montre cela en évaluant g'_2 en 3 et en $\frac{31}{10}$ en montrant la monotonie de g'_2 . On a donc *a fortiori*

$$\forall x \in \left[3, \frac{31}{10}\right], \quad 5 \leq g'_2(x) \leq 10. \quad (3.2)$$

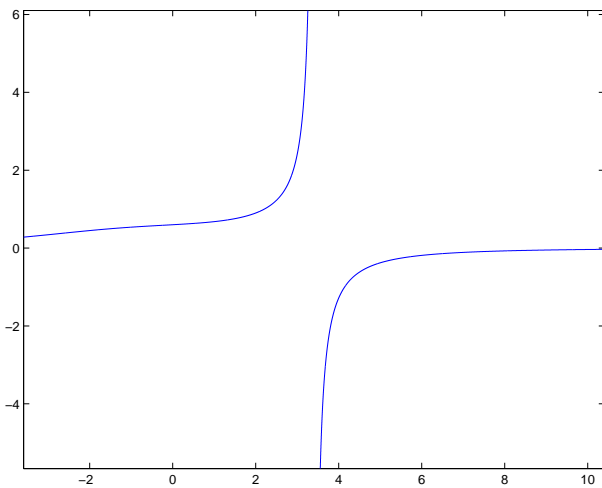
- Enfin, en posant $I = [3, \frac{31}{10}]$, on constate que si x n'appartient pas à I , g n'est pas définie si $x = \beta$ ou si g est défini, par croissance de stricte de g , on a soit

$$g(x) < g(3) = 2.4000000000000 < 3,$$

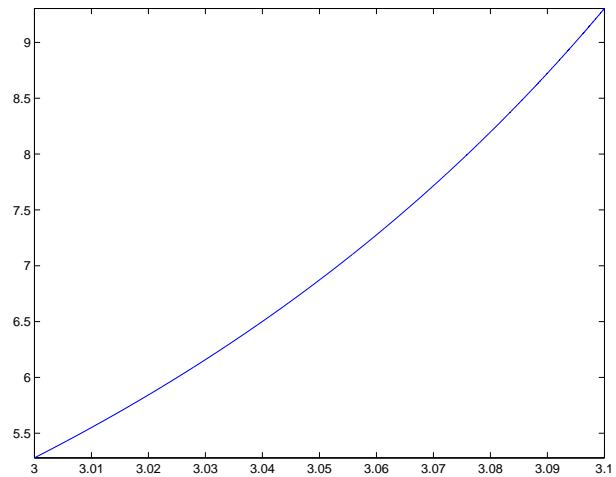
1. On remarquera que la fonction f' n'a qu'une seule racine réelle qui est 2.064960.
 2. Obtenue sous matlab symbolique.

x	$-\infty$	β	$+\infty$
Signe de $g'_3(x)$	+		+
Variations de g_3	$0^+ \nearrow +\infty$		$-\infty \nearrow 0^-$

TABLE 3.3. Tableau de variation de g_2



(a) g_2



(b) g'_2

FIGURE 3.1. Les graphes de g_2 et g'_2 (sur $[3, \frac{31}{10}]$).

soit

$$g(x) > g\left(\frac{31}{10}\right) = 3.1010422947713 > \frac{31}{10},$$

ce qui implique donc que $g(x)$ n'appartient pas à I .

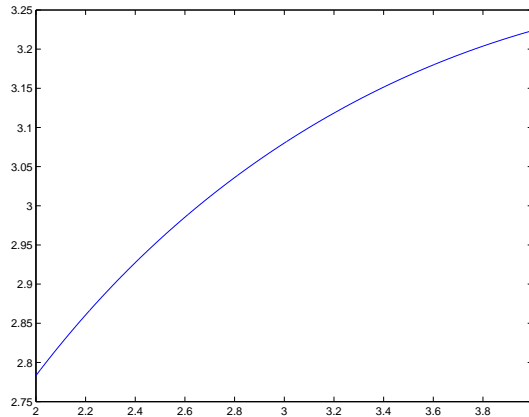
- Bref, les deux hypothèse de la proposition A.17 sont vérifiées. On en déduit la divergence de la suite du point fixe pour tout x_0 dans $I \setminus \{r\}$, où r est le point fixe de g .

(c) (i)

Voir sur le graphique 3.2, la fonction g_3 .

— D'après la question 1, f n'a qu'une solution sur $[2, 4]$. On a

$$g'_3(x) = -3 \frac{x - 5}{(-6x^2 + 60x - 36)^{3/4}}. \tag{3.3}$$

FIGURE 3.2. Le graphique de la fonction g_3 sur $[2, 4]$.

Au dénominateur de g'_3 , intervient le polynôme du second degré

$$D(x) = -6x^2 + 60x - 36,$$

dont la dérivée est nulle en 5. Ainsi ce dénominateur est strictement croissant sur $[2, 4]$. De plus, les racines de D sont 0.641101 et 9.358899. Ainsi, D est positif sur $[2, 4]$. Par ailleurs, puisque le numérateur de g'_3 est décroissant et positif, on en déduit que g'_3 est décroissante et positive sur $[2, 4]$. On déduit donc que g'_3 est majorée par son maximum qui vaut $g'_3(2) = 0.417474$ et donc le maximum de $|g'_3|$ est

$$M = 0.417474.$$

REMARQUE 3.1. Si utilise la fonction `maxabsfun`, fournie sur le site habituel, pour déterminer avec plus de soin, un majorant exact de g'_3 , définie par (3.3), on obtient

$$M = 0.417473653.$$

— On peut aussi déduire du point précédent que g_3 est strictement croissante sur $[2, 4]$. Puisque $g_3(2) = 2.783158$ et $g_3(4) = 3.223710$, la fonction g_3 laisse $I = [2, 4]$ invariant.

Des deux points précédents, on déduit, grâce à la proposition A.19, que la méthode du point fixe pour la fonction g_2 sur $[2, 4]$ est convergente vers l'unique solution recherchée sur $[2, 4]$.

(ii)

Les résultats sont donnés dans le tableau 3.4.

Voir aussi la figure 3.3 page suivante. Le tableau 3.4 et la figure 3.3 ont été faits grâce à la fonction `fixepoint`, disponible sur le site et que vous utiliserez en TP.

(3) Voir [BM03, Correction de l'exercice 4.7 p. 305].

CORRECTION DE L'EXERCICE 3.3.

(1) (a) (i) On a

$$g'(x) = -1/3 \sin(x). \quad (3.4)$$

(ii) Sur la figure 4(a), on constate que la fonction g semble avoir un point fixe, correspondant à la valeur

$$r = 0.31675082877122117188679618061096.$$

n	x_n
0	2.00000000
1	2.78315768
2	3.03201767
3	3.08657150
4	3.09731658
5	3.09938449
6	3.09978065
7	3.09985648

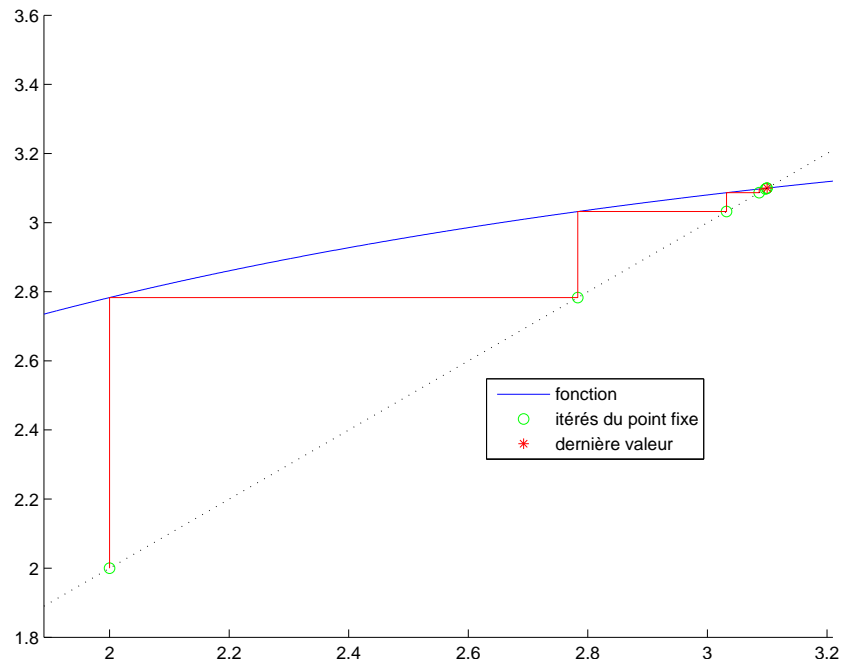
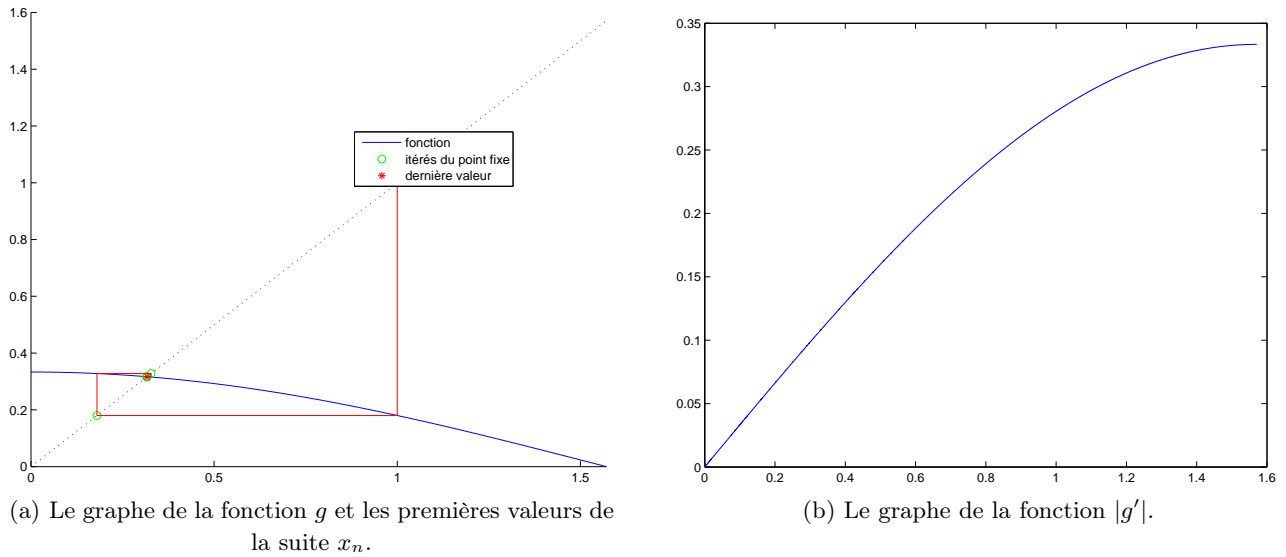
TABLE 3.4. Itérés x_n du point fixe

FIGURE 3.3. Le graphique illustrant les itérés du point fixe du tableau 3.4.

- (iii) Sur la figure 4(b), on constate que les valeurs de la fonction $|g'|$ sont inférieures à 0.333333. Démontrons cela rigoureusement. On majore la valeur absolue du sin par 1 et on majore donc la valeur absolue de la dérivée par $1/3$.
- (iv) Sur la figure 4(a), on constate que l'intervalle $[a, b]$ est g -stable. Démontrons cela rigoureusement. La fonction g est monotone (car décroissante); ainsi, sur l'intervalle $[a, b]$, elle prend les valeurs comprises entre $g(a)$ et $g(b)$. On vérifie que $g(a) = 0.333333333333$ et $g(b) = 0$ sont bien dans l'intervalle $[a, b]$.
- (b) D'après les points 1(a)iii et 1(a)iv, les deux hypothèses de la proposition A.19 sont vérifiées et donc g admet un point fixe unique r dans $I = [a, b]$ et, pour tout x_0 de I , la suite (x_n) est définie et converge vers r . Cette valeur est nécessairement celle donnée dans l'énoncé, par unicité de celle-ci !

FIGURE 3.4. Les graphes des fonctions g et $|g'|$.

- (2) Appliquons le résultat de la proposition A.21 ; on choisit n défini par (A.35), où la valeur de k a été donnée plus haut, ce qui donne numériquement

$$n = 7. \quad (3.5)$$

- (3) On obtient alors progressivement :

$$\begin{aligned} x_0 &= 1 ; \\ x_1 &= g(x_0) = 0.1801007686227 ; \\ x_2 &= g(x_1) = 0.3279418824098 ; \\ x_3 &= g(x_2) = 0.3155690860231 ; \\ x_4 &= g(x_3) = 0.3168733042460 ; \\ x_5 &= g(x_4) = 0.3167381101476 ; \\ x_6 &= g(x_5) = 0.3167521492808 ; \\ x_7 &= g(x_6) = 0.3167506916665. \end{aligned}$$

REMARQUE 3.2. Si on calcule l'erreur réellement commise, en utilisant la valeur de x_n déterminée ci-dessous et la valeur de r donnée dans l'énoncé, on a

$$|x_n - r| = |0.3167506916665 - 0.31675082877122117188679618061096| = 0.0000001371047,$$

ce qui est bien inférieur à la valeur de ε donnée dans l'énoncé.

REMARQUE 3.3. Si on utilise la majoration donnée par (L.13), on obtient

$$|x_n - r| \leq 0.0000007288071,$$

qui est bien inférieur à la valeur de ε donnée dans l'énoncé.

CORRECTION DE L'EXERCICE 3.4.

- (1) La dérivée de la fonction f est donnée par $f'(x) = -\sin(1/4 \pi x) \pi - 2e^{-x}$, que l'on écrit

$$f'(x) = -(2e^{-x} + \pi \sin(\pi x/4))$$

Sur $[0, 4]$, $\sin(\pi x/4) \geq 0$ et donc f' est strictement négative. Ainsi f est strictement décroissante. Puisque $f(0) = 6 > 0$ et que $f(4) = -4.0 < 0$, f n'a qu'un zéro noté x^* sur $[0, 4]$.

- (2) Voir la proposition A.30. On y a vu que la méthode de Newton est quadratique ssi $f'(x^*) \neq 0$ et si $f''(x^*) \neq 0$. On sait que $f'(x^*)$ est non nul, (d'après la question 1 puisque $f' < 0$), et le cours assure que $g'(x^*) = 0$. Ainsi, la méthode est au moins quadratique.

Rappelons que ce résultat est fondé sur le calcul suivant : On a, là où $f' \neq 0$,

$$g' = 1 - \left(\frac{f}{f'}\right)' = 1 - \frac{f'f' - f''f}{f'^2} = \frac{1}{f'^2} (f'^2 - f''f),$$

et donc

$$g' = \frac{f''f}{f'^2},$$

et en particulier, puisque $f'(x^*) \neq 0$,

$$g'(x^*) = \frac{f''(x^*)f(x^*)}{f'^2(x^*)}.$$

Ainsi, $g'(r) = 0$ si et seulement si $f(r) = 0$, ce qui est vrai.

De plus, elle est exactement quadratique si et seulement si $g''(x^*) \neq 0$, ce qui est équivalent à $f''(x^*) \neq 0$.

Rappelons que ce résultat est fondé sur le calcul suivant : On peut montrer que

$$g'' = \frac{1}{(f')^4} \left((f')^3 f'' + f f''' (f')^2 - f f'' ((f')^2)' \right),$$

En particulier, en x^* , puisque $f'(x^*)$ est non nul et que $f(x^*)$ est nul, on a

$$g''(x^*) = \frac{1}{(f'(x^*))^4} \left((f'(x^*))^3 f''(x^*) \right) = \frac{f''(x^*)}{f'(x^*)},$$

et donc

$$g''(x^*) = \frac{f''(x^*)}{f'(x^*)}.$$

Ainsi, $g''(x^*) \neq 0$ si et seulement si $f''(x^*) \neq 0$.

Nous avons plusieurs niveau de réponse :

- (a) Puisque x^* n'est pas (encore!) connu, nous ne connaissons pas $f''(x^*)$ et on peut affirmer seulement que la méthode de Newton est ici au moins quadratique.
- (b) Numériquement, on peut montrer (en calculant *a posteriori*³ et de façon numérique la racine x^*) que

$$f''(x^*) \approx 0.404142257991627,$$

et donc

$$f''(x^*) \neq 0. \tag{3.6}$$

Ainsi, la méthode de Newton est ici quadratique.

3. ce qui est un peu hypocrite!

(c) Enfin, à un niveau plus évolué et plus précis, on peut remarquer que l'on a

$$\begin{aligned} f''(x) &= -1/4 \cos(1/4 \pi x) \pi^2 + 2 e^{-x}, \\ f'''(x) &= 1/16 \sin(1/4 \pi x) \pi^3 - 2 e^{-x} \end{aligned}$$

et, après calculs :

$$\begin{aligned} f^{(4)}(x) &= \frac{1}{64} \cos(1/4 \pi x) \pi^4 + 2 e^{-x}, \\ f^{(5)}(x) &= -\frac{1}{256} \sin(1/4 \pi x) \pi^5 - 2 e^{-x}. \end{aligned}$$

Puisque x appartient à $[0, 4]$, on a $\pi x/4$ appartient à $[0, \pi]$; ainsi $\sin(1/4 \pi x) \geq 0$. On a aussi $e^{-x} \geq 0$ et ainsi $f^{(5)}$ est strictement négative sur $[0, 4]$. Ainsi,

$$f^{(4)} \text{ est strictement décroissante sur } [0, 4]. \quad (3.7)$$

Définissons A et B appartenant à $[0, 4]$ par

$$A = 2,$$

$$B = 2.1.$$

On vérifie que

$$f(A) \approx 0.27067057, \quad (3.8a)$$

$$f(B) \approx -0.06892353. \quad (3.8b)$$

A fortiori, d'après (3.7), on a

$$f^{(4)} \text{ est strictement décroissante sur } [A, B]. \quad (3.9)$$

On vérifie que

$$f^{(4)}(A) \approx 0.27067057 > 0, \quad (3.10a)$$

$$f^{(4)}(B) \approx 0.12549678 > 0. \quad (3.10b)$$

D'après (3.9) et (3.10), $f^{(4)}$ est strictement positive sur $[A, B]$; ainsi, f''' est strictement croissante sur $[A, B]$. Ainsi,

$$\forall x \in [A, B], \quad f'''(x) \geq f'''(A) \approx 0.27067057 > 0.$$

Ainsi,

$$f'' \text{ est strictement croissante sur } [A, B]. \quad (3.11)$$

D'après (3.8),

$$\text{l'unique zéro } x^* \text{ de } f \text{ sur } [0, 4] \text{ appartient à } [A, B]. \quad (3.12)$$

On a enfin

$$f''(A) \approx 0.27067057 > 0. \quad (3.13)$$

Ainsi, (3.11), (3.12) et (3.13) impliquent

$$f''(x^*) > 0.$$

ce qui confirme (3.6).

- (d) Une dernière méthode, beaucoup plus rapide et élégante, proposée par Guillaume Rousseau, 3A en 2022, est donnée. Raisonnons par l'absurde et supposons que

$$f''(x^*) = 0. \quad (3.14)$$

On rappelle que

$$f''(x^*) = -\frac{\pi^2}{4} \cos\left(\frac{\pi x^*}{4}\right) + 2e^{-x^*}. \quad (3.15)$$

On a aussi

$$f(x^*) = 0$$

et donc

$$4 \cos\left(\frac{\pi x^*}{4}\right) + 2e^{-x^*} = 0. \quad (3.16)$$

Si (3.14) a lieu, la différence des deux équations (3.15) et (3.16) implique que

$$\cos\left(\frac{\pi x^*}{4}\right) = 0$$

ce qui implique, en utilisant de nouveau (3.15),

$$e^{-x^*} = 0,$$

ce qui n'est guère possible!

- (3) Voir les propositions A.26 et A.30.

On y a vu que le développement de Taylor de la fonction g sur $[x^*, x_n]$ permet de montrer que

$$|x_{n+1} - x^*| \leq D |x_n - x^*|^2 \quad (3.17)$$

et $D = \frac{1}{2} \max_{x \in [0,4]} |g''(x)|$ en admettant d'abord *a priori* que la suite x_n converge et que les x_n sont dans l'intervalle $[0, 4]$.

Rappelons que ce résultat est fondé sur le calcul suivant : On applique la formule de Taylor-Lagrange à g avec $a = r$, $b = x_n$ à l'ordre $p = 1$: il existe $\xi_n \in]r, x_n[$ tel que

$$g(x_n) = g(r) + g'(r)(x_n - r) + \frac{1}{2}g''(\xi_n)(x_n - r)^2,$$

et on donc

$$x_{n+1} - x^* = g'(x^*)(x_n - x^*) + \frac{1}{2}g''(\xi_n)(x_n - x^*)^2,$$

et donc, puisque $g'(x^*) = 0$

$$x_{n+1} - x^* = \frac{1}{2}g''(\xi_n)(x_n - x^*)^2.$$

Si la suite x_n converge, alors, puisque $\xi_n \in [r, x_n]$, $\xi_n \in [r, x_n]$ tend vers r et puisque g'' est continue $g''(\xi_n)$ tend vers $g''(r)$ et donc

$$\lim_{n \rightarrow +\infty} \frac{|e_{n+1}|}{|e_n|^2} = \frac{1}{2}g''(r).$$

et donc

$$\exists D, \quad \forall n \in \mathbb{N}, \quad \frac{|e_{n+1}|}{|e_n|^2} \leq D.$$

- (4) On se place désormais sur l'intervalle $[A, B] = [1.8, 2.3]$ et l'étude précédente est encore valable sur cet intervalle puisque $f(A)f(B) = -0.7 < 0$

Nous donnons deux façons de faire.

- (a) Comme on a déjà fait dans l'exercice 1.3 page 7, nous allons prendre différentes valeurs de n , en partant de 0 et en le faisant croître jusqu'à ce que l'erreur soit inférieure à la valeur ε donnée par l'équation (3.10) de l'énoncé. Pour cela, on écrit simplement et successivement

$$|x_0 - x^*| \leq B - A = 2.3 - 1.8 = 0.5$$

puis, en utilisant les inégalités de (3.8) et (3.9) de l'énoncé, on écrit

$$|x_1 - x^*| \leq D|x_0 - x^*|^2 \leq D \times (5 \cdot 10^{-1})^2 = 5 \cdot 10^{-2},$$

ce qui est strictement supérieur à la valeur ε donnée par l'équation (3.10) de l'énoncé. On utilise de nouveau l'inégalité (3.8) de l'énoncé et on obtient

$$|x_2 - x^*| \leq D|x_1 - x^*|^2 \leq D \times (5 \cdot 10^{-2})^2 = 5 \cdot 10^{-4},$$

ce qui est strictement supérieur à la valeur ε donnée par l'équation (3.10) de l'énoncé. On utilise de nouveau l'inégalité (3.8) de l'énoncé et on obtient

$$|x_3 - x^*| \leq D|x_2 - x^*|^2 \leq D \times (5 \cdot 10^{-4})^2 = 5 \cdot 10^{-8},$$

ce qui est strictement supérieur à la valeur ε donnée par l'équation (3.10) de l'énoncé. Enfin, on utilise de nouveau l'inégalité (3.8) de l'énoncé et on obtient

$$|x_4 - x^*| \leq D|x_3 - x^*|^2 \leq D \times (5 \cdot 10^{-8})^2 = 5 \cdot 10^{-16},$$

ce qui cette fois-ci inférieure à la valeur ε donnée par l'équation (3.10) de l'énoncé. On a donc montré que l'entier n valait

$$n = 4. \tag{3.18}$$

- (b) En fait, on a écrit les différentes inéquations suivantes :

$$|x_0 - x^*| \leq 0.5 = B - A,$$

puis d'après l'inégalité (3.8) de l'énoncé :

$$|x_1 - x^*| \leq D|x_0 - x^*|^2 \leq D(B - A)^2,$$

puis d'après l'inégalité (3.8) de l'énoncé :

$$|x_2 - x^*| \leq D|x_1 - x^*|^2 \leq D(D(B - A)^2)^2,$$

puis d'après l'inégalité (3.8) de l'énoncé :

$$|x_3 - x^*| \leq D|x_2 - x^*|^2 \leq D(D(D(B - A)^2)^2)^2,$$

puis d'après l'inégalité (3.8) de l'énoncé :

$$|x_4 - x^*| \leq D|x_3 - x^*|^2 \leq D(D(D(D(B - A)^2)^2)^2)^2.$$

Autrement dit, il s'agit de calculer explicitement les différents termes de droite en fonction de n . On écrit, en notant $p = 2$ et $e_n = x_n - x^*$, tout d'abord

$$|e_0| \leq (B - A).$$

Puis, d'après l'inégalité (3.8) de l'énoncé (où p est remplacé par 2) :

$$|e_1| \leq D|e_0|^p = D(B - A)^p,$$

Puis, de nouveau d'après l'inégalité (3.8) de l'énoncé

$$|e_2| \leq D|e_1|^p = D(D(B - A)^p)^p = D^{1+p}(B - A)^{(p^2)},$$

Puis, de nouveau d'après l'inégalité (3.8) de l'énoncé

$$|e_3| \leq D|e_2|^p = D\left(D^{1+p}(B - A)^{p^2}\right)^p = D^{1+p+p^2}(B - A)^{(p^3)}.$$

Par une récurrence immédiate, on montre alors que

$$\forall n \in \mathbb{N}, \quad |e_n| \leq D^{1+p+p^2+\dots+p^{n-1}} (B-A)^{(p^n)}. \quad (3.19)$$

Puisque $p \neq 1$ on a

$$1 + p + p^2 + \dots + p^{n-1} = \frac{p^n - 1}{p - 1}$$

et donc, on réécrit (3.19) sous la forme

$$\forall n \in \mathbb{N}, \quad |x_n - x^*| \leq D^{\left(\frac{p^n - 1}{p - 1}\right)} (B-A)^{(p^n)}. \quad (3.20)$$

soit encore

$$\begin{aligned} |x_n - x^*| &= \leq D^{\left(\frac{p^n - 1}{p - 1}\right)} (B-A)^{(p^n)}, \\ &= \leq D^{\left(\frac{-1}{p-1}\right)} D^{\left(\frac{p^n}{p-1}\right)} (B-A)^{(p^n)}, \\ &= \leq D^{\left(\frac{1}{1-p}\right)} \left(D^{\left(\frac{1}{p-1}\right)}\right)^{(p^n)} (B-A)^{(p^n)}, \\ &= \leq D^{\left(\frac{1}{1-p}\right)} \left((B-A)D^{\left(\frac{1}{p-1}\right)}\right)^{(p^n)}, \end{aligned}$$

soit, en notant

$$\gamma = D^{\left(\frac{1}{1-p}\right)}, \quad (3.21a)$$

$$\delta = (B-A)D^{\left(\frac{1}{p-1}\right)}, \quad (3.21b)$$

on a donc

$$\forall n \in \mathbb{N}, \quad |x_n - x^*| \leq \gamma \delta^{(p^n)}. \quad (3.22)$$

On vérifie numériquement que δ , donné par

$$\delta = 0.1, \quad (3.23)$$

vérifie

$$\delta < 1.$$

Ainsi, la limite de $|x_n - x^*|$ est bien nulle quand n tend vers l'infini! Cela justifie *a posteriori* l'hypothèse faite dans la question 3, à savoir que la suite x_n converge et que tous les x_n sont bien dans l'intervalle $[0, 4]$. Voir en effet la proposition A.26.

Pour avoir

$$|x_n - x^*| \leq \varepsilon$$

il suffit donc que

$$\gamma \delta^{(p^n)} \leq \varepsilon,$$

ce qui est équivalent à

$$\delta^{(p^n)} \leq \frac{\varepsilon}{\gamma}.$$

Si on prend le logarithme de cette inégalité, on arrive à

$$p^n \geq \frac{\ln \frac{\varepsilon}{\gamma}}{\ln \delta}.$$

Si on prend de nouveau le logarithme de cette inégalité, on arrive à

$$n \ln p \geq \ln \left(\frac{\ln \frac{\varepsilon}{\gamma}}{\ln \delta} \right).$$

et donc

$$n \geq \frac{1}{\ln p} \ln \left(\frac{\ln \frac{\varepsilon}{\gamma}}{\ln \delta} \right). \quad (3.24)$$

Numériquement, on a

$$p = 2,$$

la valeur de δ est donnée par (3.23) et, d'après (3.21a),

$$\gamma = 5,$$

et donc

$$n \geq 4. \tag{3.25}$$

et on retrouve donc (3.18).

REMARQUE 3.4. On vient en fait de démontrer la proposition A.24 dans le cas $p = 2$.

REMARQUE 3.5. Si on fait tourner effectivement la méthode de Newton en prenant $x_0 = 2.3$, on arrive à

$$x_4 = 2.079615501859806$$

Si on se sert de la fonction `fzero` de matlab, on arrive à

$$x^* = 2.079615501859807$$

dont on peut supposer que c'est le zéro exact de f . La différence entre ces deux derniers nombres vaut $4.44 \cdot 10^{-16}$ et est bien inférieure à 10^{-10} .

CORRECTION DE L'EXERCICE 3.5.

(1) (a) On constate dans le tableau de l'énoncé que la suite des itérés semble converger vers

$$\tilde{x} = 1.0000000000000000.$$

(b) On constate aussi que le nombre de chiffres significatifs exacts est approximativement doublé à chaque étape, ce qui caractérise l'ordre 2 de la méthode (voir [BM03, Corollaire D.4 p. 367]).

(c)

On peut confirmer cela en traçant le graphique du nuage de points $(|e_n|, |e_{n+1}|)$, comme en figure 3.5. Les points y sont très bien alignés. On sait d'après le cours qu'une méthode d'ordre q vérifie

$$\lim_{n \rightarrow +\infty} \frac{|e_{n+1}|}{|e_n|^q} = C,$$

(où $e_n = x_n - x^*$), ce qui implique en prenant le logarithme en base 10 que l'on a asymptotiquement

$$\log(|e_{n+1}|) \approx q \log(|e_n|) + K.$$

la pente du nuage fournit donc q . Numériquement, on a une corrélation égale à 0.9993946 et une pente égale à 1.9269961, ce qui confirme numériquement l'ordre 2 de la méthode de Newton.

(2) (a) Si on regarde le tableau 3.2 de l'énoncé, on constate cette fois que la convergence a l'air beaucoup plus lente, ce qui est aussi confirmé par l'étude du graphique 3.4 de l'énoncé. La suite des itérés semble converger vers

$$\tilde{x} = 2.999991131057611.$$

(b) Le nombre de chiffres significatifs exacts ne semble plus approximativement doublé à chaque étape,

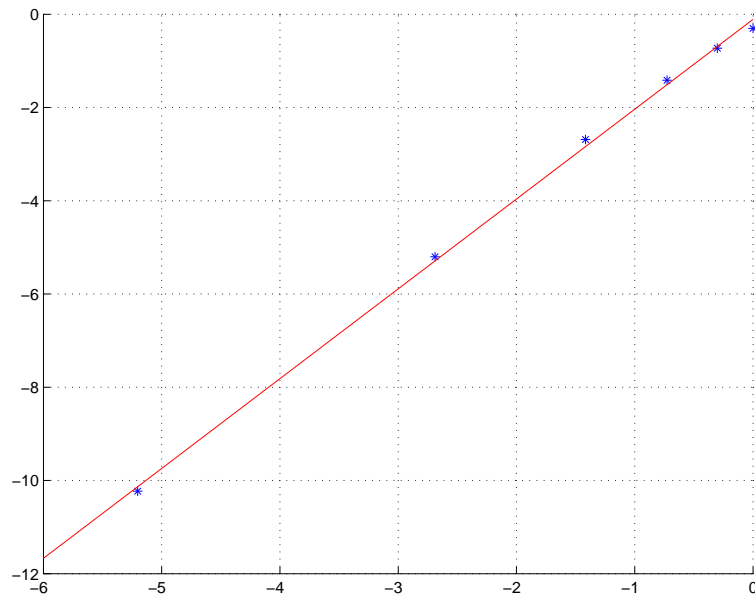


FIGURE 3.5. Nuage des points et la droite aux moindres carrés.

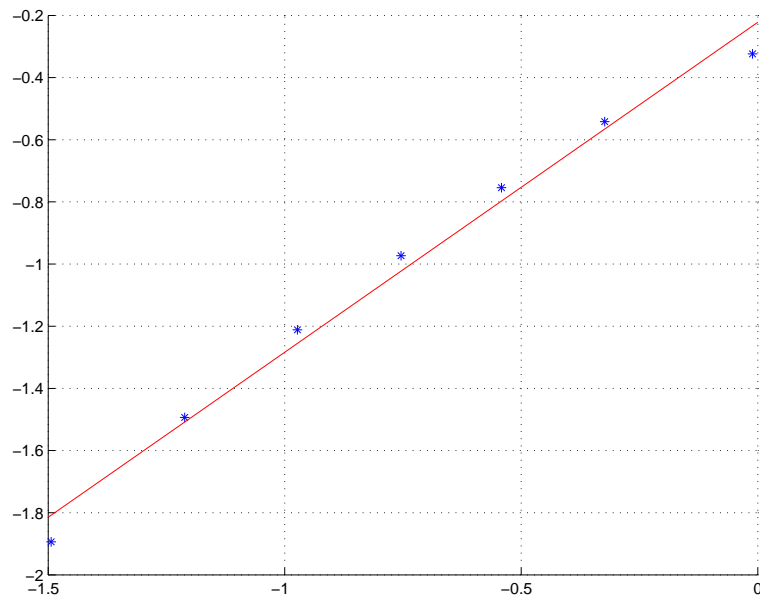


FIGURE 3.6. Nuage des points et la droite aux moindres carrés.

- (c) Si on raisonne comme dans la question 1c, on obtient le graphique de la figure 3.6. Numériquement, on a une corrélation égale à 0.9937810 et une pente égale à 1.0617333, ce qui confirme numériquement l'ordre 1 de la méthode de Newton.

Dans le cours, il est écrit que la méthode de Newton appliquée à la résolution de $p(x) = 0$ en x^* est d'ordre 2 si $p'(x^*) \neq 0$, ce qui est le cas de la question 1. Cela est confirmé par le graphique de l'énoncé et le tableau 3.5.

n	$f'(x_n)$
0	-54.000000000000000
1	-25.000000000000000
2	-13.159179687500000
3	-8.947156337219283
4	-8.049255577028404
5	-8.000150550479987
6	-8.000000001416552
7	-8.000000000000000
8	-8.000000000000000

TABLE 3.5. Valeurs des $f'(x_n)$

n	$f'(x_n)$
0	2.000000000000000
1	1.000000000000000
2	0.463867187500000
3	0.210849980533538
4	0.094952651329073
5	0.042545238988474
6	0.019006769239766
7	0.008475659295463
8	0.003775173363366
37	0.000000004584606
38	0.000000003614076
39	0.000000002311488
40	0.000000000471942
41	0.000000000471942

TABLE 3.6. Valeurs des $f'(x_n)$

En revanche, dans le cas de la question 2, il semblerait que f' soit aussi nulle en x^* , comme le montrent la figure de l'énoncé et le tableau 3.6. Nous sommes donc dans le cas où l'ordre de la méthode de Newton n'est plus 2. On peut montrer en théorie que si x^* est une racine double de f , alors la méthode de Newton est dégénérée d'ordre 1. Plus de détails dans l'annexe J.

REMARQUE 3.6. Reprenons les simulations faites dans la question 2 en utilisant la méthode de Newton modifiée, présentée dans l'annexe J. Elle consiste à écrire

$$x_{n+1} = x_n - A \frac{p(x_n)}{p'(x_n)}, \quad (3.26)$$

où A est égal à l'ordre de multiplicité recherché, ici égal à $m = 3$. On peut en effet montrer que

$$p(x) = x^4 - 10x^3 + 36x^2 - 54x + 27 = (x - 1)(x - 3)^3. \quad (3.27)$$

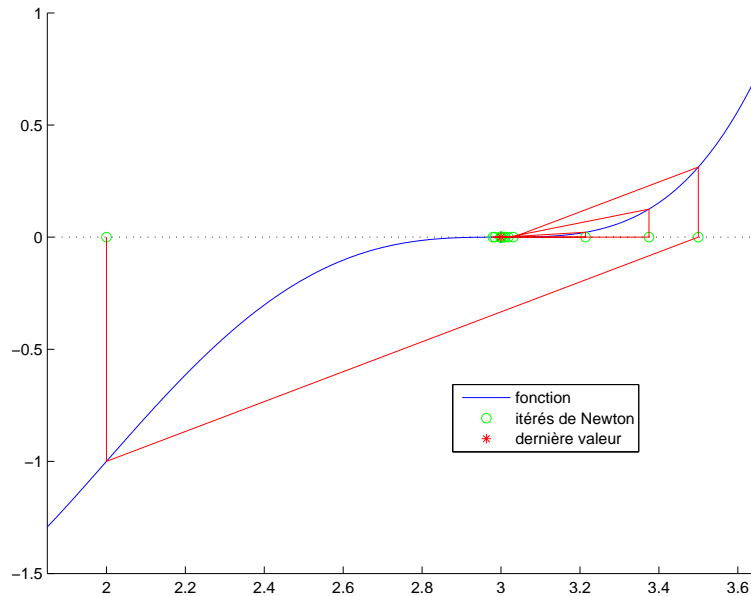


FIGURE 3.7. Graphiques des itérés de la méthode de Newton modifiée pour $x_0 = 2$.

L'application de cette méthode de Newton modifiée à l'équation (3.12) des sujets de TD a produit les résultats donnés dans les tableaux 3.7 page suivante et sur la figure 3.7.

Constatons tout d'abord que cette méthode de Newton modifiée donne bien des nombres d'itérations plus faible que la méthode de Newton habituelle (comparer les tableaux 3.7 page suivante avec les tableaux 3.2 de l'énoncé). On constate aussi que pour $n = 28$, on a numériquement

$$\begin{aligned} x_{n-1} &= 2.999814922991660, \\ x_n &= 2.999999979230154, \\ p(x_{n-1}) &= -1.26760824059602 \cdot 10^{-11}, \\ p(x_n) &= 0, \end{aligned}$$

ce qui explique l'arrêt de l'algorithme pour cette valeur de n .

La détermination graphique de l'ordre a donné la figure 3.8 page 49.

Numériquement, on a une corrélation égale à 1.0000000 et une pente égale à 1.9627567, ce qui confirme numériquement l'ordre de nouveau égal à 2 de la méthode de Newton modifiée.

En fait ce calcul de pente s'est fait uniquement sur les dernières valeurs des suites obtenues. Une observation précise des valeurs obtenues montre que le nombre de chiffres exacts ne double pas exactement globalement à chaque itération, comme théoriquement attendu. Cela provient d'arrondis de calcul. Voir la remarque J.2 page 136.

n	x_n
0	2.0000000000000000
1	3.5000000000000000
2	3.0312500000000000
3	3.000159438775510
4	3.000000412667420
5	2.979310757495006
6	3.000072338539754
7	3.000002073328751
8	2.999175854656209
9	3.000000113388043
10	3.375000113388043
11	3.018750010204911
12	3.000057870426068
13	3.000002709210343
14	3.001697350914668
15	3.000000479596380
16	2.984456438145603
17	3.000040688668089
18	3.000000990289929
19	3.010910081199020
20	3.000019695010451
21	2.999992218624878
22	3.000226923678860
23	3.000000111316675
24	3.214285825602390
25	3.006696435093944
26	3.000007440478665
27	2.999814922991660
28	2.99999979230154

TABLE 3.7. Itérations de la méthode de Newton modifiée pour $x_0 = 2$.

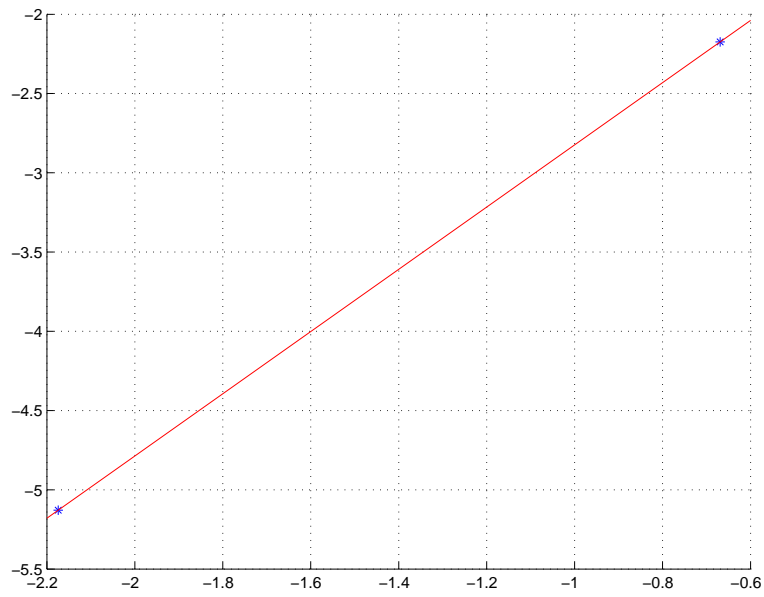


FIGURE 3.8. Nuage des derniers points et la droite aux moindres carrés (méthode de Newton modifiée).

CORRECTION DU TRAVAUX DIRIGÉS 4

Équations différentielles

CORRECTION DE L'EXERCICE 4.1.

(1) En posant $\xi_0 = 3$ et

$$f(t, y) = t^3 y^2, \quad (4.1)$$

l'équation différentielle

$$\forall t \in [0, T], \quad y'(t) = t^3 y(t)^2, \quad (4.2a)$$

$$y(0) = 3, \quad (4.2b)$$

est équivalente à

$$\forall t \in [0, T], \quad y'(t) = f(t, y(t)), \quad (4.3a)$$

$$y(0) = \xi_0. \quad (4.3b)$$

On calcule pour $n \in \{1, \dots, N = 3\}$, les approximations $y_n \approx y(t_n)$ en utilisant les définitions A.34, A.36 et A.37.

n	y_n
0	3.00000000
1	3.00000000
2	3.00090000
3	3.00810432

TABLE 4.1. Solutions approchées avec Euler explicite

n	y_n
0	3.00000000
1	3.00045000
2	3.00450338
3	3.02035935

TABLE 4.2. Solutions approchées avec RK2

Les résultats sont donnés dans les tableaux 4.1, 4.2 et 4.3.

n	y_n
0	3.00000000
1	3.00022501
2	3.00360428
3	3.01833625

TABLE 4.3. Solutions approchées avec RK4

n	$ y(t_n) - y_n $
0	0.00000000
1	0.00022502
2	0.00270433
3	0.01023207

TABLE 4.4. Erreurs avec Euler explicite

n	$ y(t_n) - y_n $
0	0
1	$2.249831236 \cdot 10^{-4}$
2	$8.990511972 \cdot 10^{-4}$
3	$2.22957175 \cdot 10^{-3}$

TABLE 4.5. Erreurs avec RK2

n	$ y(t_n) - y_n $
0	0
1	$4.219370187 \cdot 10^{-9}$
2	$4.219804994 \cdot 10^{-8}$
3	$1.438824624 \cdot 10^{-7}$

TABLE 4.6. Erreurs avec RK4

(2)

Les résultats en erreur sont donnés dans les tableaux 4.4, 4.5 et 4.6. On y constate que l'erreur est de plus en plus faible selon la méthode utilisée.

(3) On peut proposer $y'(t_n) \approx (y_{n+1} - y_n)/(t_n)$, puisque $y_n \approx y(t_n)$. Cependant, dans cette expression, on constate qu'on a deux approximations différentes de la dérivées entrant en jeux. Rien dans le cours ne justifie cette approximation. Cependant puisque f , définie par (4.1), est continue en second argument,

que $y'(t_n) = f(t_n, y(t_n))$ et que $y_n \rightarrow y(t_n)$, on peut alors proposer

$$y'(t_n) \approx f(t_n, y_n).$$

CORRECTION DE L'EXERCICE 4.2.

En posant $\xi_0 = 1$ et

$$f(t, y) = -t^2 + yt, \quad (4.4)$$

l'équation différentielle

$$\forall t \in [0, T], \quad y'(t) = -t^2 + y(t)t, \quad (4.5a)$$

$$y(0) = 1, \quad (4.5b)$$

est équivalente à

$$\forall t \in [0, T], \quad y'(t) = f(t, y(t)), \quad (4.6a)$$

$$y(0) = \xi_0. \quad (4.6b)$$

On calcule pour $n \in \{1, \dots, N = 2\}$, les approximations $y_n \approx y(t_n)$.

n	y_n
0	1.00000000
1	1.00000000
2	1.06300000

TABLE 4.7. Solutions approchées avec Euler explicite

n	y_n
0	1.00000000
1	1.03150000
2	1.10917765

TABLE 4.8. Solutions approchées avec RK2

n	y_n
0	1.00000000
1	1.03687353
2	1.11977332

TABLE 4.9. Solutions approchées avec RK4

Les résultats sont donnés dans les tableaux 4.7, 4.8 et 4.9, obtenus en utilisant les définitions A.34, A.36 et A.37.

CORRECTION DE L'EXERCICE 4.3.

On pose

$$\begin{aligned}y_1 &= y, \\y_2 &= y', \\y_3 &= y''.\end{aligned}$$

On a alors, pour tout t ,

$$\begin{aligned}y_1'(t) &= y'(t) = y_2(t), \\y_2'(t) &= y''(t) = y_3(t), \\y_3'(t) &= y^{(3)}(t) = y''(t) - 2y'(t) + y(t) - 2 = y_3(t) - 2y_2(t) + y_1 - 2.\end{aligned}$$

Ainsi, en posant

$$Y(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix},$$

on a

$$Y'(t) = \begin{pmatrix} y_2(t) \\ y_3(t) \\ y_3(t) - 2y_2(t) + y_1(t) - 2 \end{pmatrix},$$

On obtient alors

$$\forall t \in [0, T], \quad Y'(t) = F(t, Y(t)), \quad (4.7a)$$

$$Y(0) = \Xi_0, \quad (4.7b)$$

où

$$\forall t \in [0, T], \quad \forall Y = {}^t(y_1, y_2, y_3) \in \mathbb{R}^3, \quad F(t, Y) = \begin{pmatrix} y_2 \\ y_3 \\ y_3 - 2y_2 + y_1 - 2 \end{pmatrix}, \quad (4.8a)$$

$$\Xi_0 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}. \quad (4.8b)$$

Remarquons qu'ici, le système est affine puisque

$$F(t, Y) = AY + B,$$

où

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -2 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0 \\ -2 \end{pmatrix}$$

REMARQUE 4.1. Cette opération est essentielle aussi bien sur le plan théorique que numérique. Par exemple, sous matlab, il faut définir la fonction F avant d'utiliser un solveur. Ici, il faudrait par exemple taper

```
f=inline(' [0 1 0; 0 0 1; 1 -2 1]*Y+[0;0;2]', 't', 'Y');
Y0=[0;1;2];
[T,Y]=ode45(f,[0 1],Y0);
plot(T,Y);
```

ce qui fournirait le graphique 4.1.

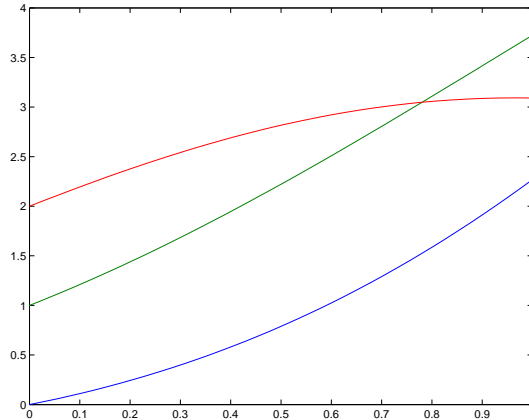


FIGURE 4.1. Le graphique de résolution de (4.7)

CORRECTION DE L'EXERCICE 4.4.

Voir aussi [BM03, Correction de l'exercice 5.7 p. 323].

On pose

$$y_1 = x_1, \quad y_2 = \dot{x}_1, \quad y_3 = x_2, \quad y_4 = \dot{x}_2$$

Ainsi, les équations différentielles de l'énoncé sont équivalentes au système

$$\begin{aligned} \dot{y}_1(t) &= y_2(t), \\ \dot{y}_2(t) &= -(k_1 + k_2) y_1(t) + k_2 y_3(t) - c_1 y_2(t) - d_1 y_2^3(t) + f_1(t), \\ \dot{y}_3(t) &= y_4(t), \\ \dot{y}_4(t) &= k_2 y_1(t) - (k_2 + k_3) y_3(t) - c_2 y_4(t) - d_2 y_4^3(t) + f_2(t). \end{aligned}$$

Ainsi, si l'on se donne les conditions initiales sous la forme

$$\Xi_0 = \begin{pmatrix} x_1(0) \\ \dot{x}_1(0) \\ x_2(0) \\ \dot{x}_2(0) \end{pmatrix}$$

alors les équations différentielles de l'énoncé le système s'écrivent sous la forme (4.6) de l'énoncé avec $p = 4$.

La fonction F associée au système des équations différentielles de l'énoncé est donnée par

$$F \left(t, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} \right) = \begin{pmatrix} y_2 \\ -(k_1 + k_2) y_1 + k_2 y_3 - c_1 y_2 - d_1 y_2^3 + f_1(t), \\ y_4 \\ k_2 y_1 - (k_2 + k_3) y_3 - c_2 y_4 - d_2 y_4^3 + f_2(t) \end{pmatrix},$$

et le vecteur des conditions initiales est donné par

$$\Xi_0 = \begin{pmatrix} x_1(0) \\ \dot{x}_1(0) \\ x_2(0) \\ \dot{x}_2(0) \end{pmatrix}$$

Pour une utilisation informatique, on peut aussi mettre cela sous la forme

On considère l'application G de $[0, T]$ dans \mathbb{R}^4 définie par

$$\forall t \in [a, b], \quad G(t) = \begin{pmatrix} 0 \\ f_1(t) \\ 0 \\ f_2(t) \end{pmatrix},$$

l'application H de \mathbb{R}^4 dans \mathbb{R}^4 définie par

$$\forall Y \in \mathbb{R}^4, \quad H(Y) = H \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = - \begin{pmatrix} 0 \\ d_1 y_2^3 \\ 0 \\ d_2 y_4^3 \end{pmatrix}.$$

Enfin, on considère A la matrice $\mathcal{M}_4(\mathbb{R})$ définie par

$$A = - \begin{pmatrix} 0 & -1 & 0 & 0 \\ k_1 + k_2 & c_1 & -k_2 & 0 \\ 0 & 0 & 0 & -1 \\ -k_2 & 0 & k_2 + k_3 & c_2 \end{pmatrix}.$$

On définit enfin l'application F de $[0, T] \times \mathbb{R}^4$ dans \mathbb{R}^4 par

$$\forall t \in [a, b], \quad \forall Y \in \mathbb{R}^4, \quad F(t, Y) = G(t) + H(Y) + AY.$$

◇

n	$y_{1,n}$	$y_{2,n}$	$y_{3,n}$	$y_{4,n}$
0	1.00000000	1.00000000	2.00000000	1.00000000
1	1.01000000	0.99980000	2.01000000	0.96980000
2	1.01999800	0.99950008	2.01969800	0.93951181
3	1.02999300	0.99909730	2.02909312	0.90914095
4	1.03998397	0.99858873	2.03818453	0.87869296
5	1.04996986	0.99797146	2.04697146	0.84817339

TABLE 4.10. Solutions approchées avec Euler

Les résultats des simulations sont donnés dans les tableaux 4.10.

On a $Y_0 = \Xi_0$ où

$$\Xi_0 = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \end{pmatrix},$$

ce qui fournit la première ligne du tableau 4.10. Précisons maintenant le passage de Y_0 à Y_1 . D'après le schéma donné dans la définition A.34 page 73 appliqué à F et Y , pour $n = 0$, on a

$$Y_1 = Y_0 + hF(t_0, Y_0).$$

On a aussi

$$F(0, Y_0) = \begin{pmatrix} y_{2,0} \\ -(k_1 + k_2)y_{1,0} + k_2 y_{3,0} - c_1 y_{2,0} - d_1 y_{2,0}^3 + f_1(0) \\ y_{4,0} \\ k_2 y_{1,0} - (k_2 + k_3)y_{3,0} - c_2 y_{4,0} - d_2 y_{4,0}^3 + f_2(0) \end{pmatrix},$$

et donc, compte tenu des conditions initiales et des données de l'énoncé

$$F(0, Y_0) = \begin{pmatrix} 1 \\ -2 \times 1 + 1 \times 2 - 0,01 \times 1 - 0,01 \times 1^3 + 0 \\ 1 \\ 1 \times 1 - 2 \times 2 - 0,01 \times 1 - 0,01 \times 1^3 + 0 \end{pmatrix},$$

et donc

$$F(0, Y_0) = \begin{pmatrix} 1 \\ -0.02000000 \\ 1 \\ -3.02000000 \end{pmatrix},$$

et enfin

$$Y_1 = Y_0 + hF(0, Y_0) = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \end{pmatrix} + 0,01 \times \begin{pmatrix} 1 \\ -0.02000000 \\ 1 \\ -3.02000000 \end{pmatrix},$$

et donc

$$Y_1 = \begin{pmatrix} 1.01000000 \\ 0.99980000 \\ 2.01000000 \\ 0.96980000 \end{pmatrix}$$

ce qui fournit la deuxième ligne du tableau 4.10. On recommence. D'après le schéma donné dans la définition A.34 page 73 appliqué à F et Y , pour $n = 1$, on a

$$Y_2 = Y_1 + hF(t_1, Y_1),$$

où Y_1 vient d'être calculé et $t_1 = 1 \times h = h = 0,01$. Or

$$F(t_1, Y_1) = \begin{pmatrix} y_{2,1} \\ -(k_1 + k_2)y_{1,1} + k_2y_{3,1} - c_1y_{2,1} - d_1y_{2,1}^3 + f_1(0,01) \\ y_{4,1} \\ k_2y_{1,1} - (k_2 + k_3)y_{3,1} - c_2y_{4,1} - d_2y_{4,1}^3 + f_2(0,01) \end{pmatrix},$$

et donc

$$F(t_1, Y_1) = \begin{pmatrix} 0.99980000 \\ -2 \times 1.01000000 + 1 \times 2.01000000 - 0,01 \times 0.99980000 - 0,01 \times (0.99980000)^3 + 0 \\ 0.96980000 \\ 1 \times 1.01000000 - 2 \times 2.01000000 - 0,01 \times 0.96980000 - 0,01 \times 0.96980000^3 + 0 \end{pmatrix},$$

et donc

$$F(t_1, Y_1) = \begin{pmatrix} 0.99980000 \\ -0.02999200 \\ 0.96980000 \\ -3.02881909 \end{pmatrix},$$

et enfin

$$Y_2 = Y_1 + hF(t_1, Y_1) = \begin{pmatrix} 1.01000000 \\ 0.99980000 \\ 2.01000000 \\ 0.96980000 \end{pmatrix} + 0,01 \times \begin{pmatrix} 0.99980000 \\ -0.02999200 \\ 0.96980000 \\ -3.02881909 \end{pmatrix},$$

et donc

$$Y_2 = \begin{pmatrix} 1.01999800 \\ 0.99950008 \\ 2.01969800 \\ 0.93951181 \end{pmatrix}$$

ce qui fournit la troisième ligne du tableau 4.10.

Pour chaque valeur de $i \in \{0, \dots, N = 5\}$, sont données les approximations $Y_i = (y_{1,i}, y_{2,i}, y_{3,i}, y_{4,i})$ qui représentent donc des approximations de $x_1(t_i)$, $\dot{x}_1(t_i)$, $x_2(t_i)$ et $\dot{x}_2(t_i)$.

Systèmes d'équations linéaires

CORRECTION DE L'EXERCICE 5.1.

On obtient la matrice triangulaire supérieure suivante

$$M = \begin{pmatrix} 3 & 6 & 15 \\ 0 & -5 & -8 \\ 0 & 0 & -1 \end{pmatrix}$$

avec un second membre égal à

$$c = \begin{pmatrix} 24 \\ -13 \\ -1 \end{pmatrix}$$

La résolution du système triangulaire supérieur fournit donc la solution

$$X = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

- (1) Précisons la phase de transformation de la matrice en une matrice triangulaire supérieure, en précisant à chaque fois le pivot considéré, qui est encadré :

— La matrice augmentée est

$$\begin{pmatrix} 1 & -\frac{1}{2} & 0 & \frac{1}{2} \\ 2 & -1 & 2 & 3 \\ \boxed{3} & 6 & 15 & 24 \end{pmatrix}$$

— Première étape : on permute L_3 et L_1 . La nouvelle matrice augmentée est donc

$$\begin{pmatrix} \boxed{3} & 6 & 15 & 24 \\ 2 & -1 & 2 & 3 \\ 1 & -\frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

On effectue les opérations de lignes

$$\begin{aligned} L_2 &\leftarrow L_2 - \frac{2}{3}L_1, \\ L_3 &\leftarrow L_3 - \frac{1}{3}L_1. \end{aligned}$$

La nouvelle matrice augmentée est donc

$$\begin{pmatrix} 3 & 6 & 15 & 24 \\ 0 & \boxed{-5} & -8 & 3 - \frac{2}{3} \times 24 \\ 0 & -\frac{5}{2} & -5 & \frac{1}{2} - \frac{1}{3} \times 24 \end{pmatrix}$$

soit

$$\begin{pmatrix} 3 & 6 & 15 & 24 \\ 0 & \boxed{-5} & -8 & -13 \\ 0 & -\frac{5}{2} & -5 & -\frac{15}{2} \end{pmatrix}$$

— Deuxième étape : Aucune permutation de lignes n'est nécessaire. On effectue ensuite l'opération de ligne

$$L_3 \leftarrow L_3 - \frac{1}{2}L_2.$$

La nouvelle matrice est donc

$$\begin{pmatrix} 3 & 6 & 15 & 24 \\ 0 & -5 & -8 & -13 \\ 0 & 0 & \boxed{-1} & -\frac{15}{2} + \frac{1}{2} \times 13 \end{pmatrix}$$

soit

$$\begin{pmatrix} 3 & 6 & 15 & 24 \\ 0 & -5 & -8 & -13 \\ 0 & 0 & \boxed{-1} & -1 \end{pmatrix}$$

(2) On résout enfin le système triangulaire obtenu en commençant par le bas : la troisième équation fournit

$$x_3 = \frac{-1}{-1} = 1.$$

La deuxième équation fournit ensuite successivement

$$\begin{aligned} x_2 &= -\frac{1}{5}(-13 + 8x_3), \\ &= -\frac{1}{5}(-13 + 8), \\ &= 1. \end{aligned}$$

La première équation fournit enfin successivement

$$\begin{aligned} x_1 &= \frac{1}{3}(24 - 6x_2 - 15x_3), \\ &= \frac{1}{3}(24 - 6 - 15), \\ &= 1. \end{aligned}$$

CORRECTION DE L'EXERCICE 5.2.

On obtient la matrice triangulaire supérieure suivante

$$M = \begin{pmatrix} 9 & 0 & -2 \\ 0 & 8 & \frac{13}{9} \\ 0 & 0 & \frac{56}{9} \end{pmatrix}$$

avec un second membre égal à

$$c = \begin{pmatrix} 5 \\ \frac{242}{9} \\ \frac{112}{9} \end{pmatrix}$$

La résolution du système triangulaire supérieur fournit donc la solution,

$$\tilde{X} = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix},$$

où l'on a opéré la permutation des colonnes données par $\{2, 1, 3\}$. Après permutation des colonnes, on a donc la solution

$$X = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}.$$

On peut vérifier que

$$AX = \begin{pmatrix} 13 \\ 5 \\ 28 \end{pmatrix} = B.$$

Précisons la phase de transformation de la matrice en une matrice triangulaire supérieure, en précisant à chaque fois le pivot considéré, qui est encadré :

— La matrice augmentée est

$$\begin{pmatrix} 0 & 1 & 6 & 13 \\ 0 & 9 & -2 & 5 \\ 8 & 2 & 1 & 28 \end{pmatrix}$$

— Première étape : on permute L_2 et L_1 , puis C_1 et C_2 . Le tableau des indices passe de $\{1, 2, 3\}$ à $\{2, 1, 3\}$. La nouvelle matrice augmentée est donc

$$\begin{pmatrix} \boxed{9} & 0 & -2 & 5 \\ 1 & 0 & 6 & 13 \\ 2 & 8 & 1 & 28 \end{pmatrix}$$

On effectue ensuite les opérations de lignes

$$\begin{aligned} L_2 &\leftarrow L_2 - \frac{1}{9}L_1, \\ L_3 &\leftarrow L_3 - \frac{2}{9}L_1. \end{aligned}$$

La nouvelle matrice est donc

$$\begin{pmatrix} 9 & 0 & -2 & 5 \\ 0 & 0 & \frac{56}{9} & \frac{112}{9} \\ 0 & \boxed{8} & \frac{13}{9} & \frac{242}{9} \end{pmatrix}$$

— Deuxième étape : on permute L_2 et L_3 (sans simplifier la deuxième ligne!). La nouvelle matrice est donc

$$\begin{pmatrix} 9 & 0 & -2 & 5 \\ 0 & \boxed{8} & \frac{13}{9} & \frac{242}{9} \\ 0 & 0 & \frac{56}{9} & \frac{112}{9} \end{pmatrix}$$

Le calcul est maintenant terminé!

CORRECTION DE L'EXERCICE 5.3.

(1) On obtient la matrice triangulaire U supérieure suivante

$$U = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

et la matrice triangulaire L inférieure suivante

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 3 & 0 \\ 1 & 2 & 3 \end{pmatrix}$$

(2) Le système $Ax = b$ est équivalent à $LUx = b$ et donc aux deux systèmes équivalents que l'on résout successivement : $Ly = b$ et $Ux = y$. La résolution de $Ly = b$ fournit

$$y = \begin{pmatrix} 6 \\ 5 \\ 3 \end{pmatrix}$$

et la résolution de $Ux = y$ fournit finalement

$$x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

(3) On peut vérifier finalement que la solution est correcte :

$$Ax = \begin{pmatrix} 6 \\ 21 \\ 25 \end{pmatrix} = b.$$

Exercices facultatifs

Non corrigés.

Rappels théoriques

A.1. Interpolation

LEMME A.1. Soit $n \in \mathbb{N}$. Pour tout $i \in \{0, \dots, n\}$, il existe un unique polynôme l_i de degré n tel que

$$\forall j \in \{0, \dots, n\} \setminus \{i\}, \quad l_i(x_j) = 0, \quad (\text{A.1a})$$

$$l_i(x_i) = 1, \quad (\text{A.1b})$$

soit encore

$$\forall j \in \{0, \dots, n\}, \quad l_i(x_j) = \delta_{ij}, \quad (\text{A.2})$$

où δ_{ij} est le symbole de Kronecker¹. Pour tout $i \in \{0, \dots, n\}$, l_i est donné par

$$l_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j}. \quad (\text{A.4})$$

Les polynômes $(l_i)_{0 \leq i \leq n}$ sont appelés les polynômes de Lagrange, relatifs aux support $\{x_0, \dots, x_n\}$.

PROPOSITION A.2. Soit $n \in \mathbb{N}$. Il existe un unique polynôme Π_n de degré au plus n tel que

$$\forall j \in \{0, \dots, n\} \quad \Pi_n(x_j) = y_j. \quad (\text{A.5})$$

Il est donné par

$$\Pi_n(x) = \sum_{i=0}^n y_i l_i(x). \quad (\text{A.6})$$

Notons que le degré de Π_n est au plus n . Il peut être plus petit !

DÉFINITION A.3 (Tableau des différences divisées).

On définit le tableau des différences divisées de la façon suivante :

- La première colonne contient les $n + 1$ valeurs $(x_i)_{0 \leq i \leq n}$, classées dans l'ordre des indices (qui n'est pas nécessairement l'ordre des valeurs).
- La deuxième colonne contient les $n + 1$ différences divisées d'ordre 0, c'est-à-dire $(f[x_i])_{0 \leq i \leq n}$, chacune d'elles étant égale à $f(x_i)$ d'après

$$\forall i \in \{0, \dots, n\}, \quad f[x_i] = f(x_i). \quad (\text{A.7})$$

- La troisième colonne contient les n différences divisées d'ordre 1, c'est-à-dire $(f[x_i, x_{i+1}])_{0 \leq i \leq n-1}$, données par

$$\forall i \in \{0, \dots, n-1\}, \quad f[x_i, x_{i+1}] = \frac{f[x_{i+1}] - f[x_i]}{x_{i+1} - x_i}. \quad (\text{A.8})$$

1. défini par

$$\forall i, j \in \{0, \dots, n\}, \quad \delta_{ij} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j. \end{cases} \quad (\text{A.3})$$

- La quatrième colonne contient les $n-1$ différences divisées d'ordre 2, c'est-à-dire $(f[x_i, x_{i+1}, x_{i+2}])_{0 \leq i \leq n-2}$, données par

$$\forall i \in \{0, \dots, n-2\}, \quad f[x_i, x_{i+1}, x_{i+2}] = \frac{f[x_{i+1}, x_{i+2}] - f[x_i, x_{i+1}]}{x_{i+2} - x_i}. \quad (\text{A.9})$$

- De façon générale, la $k+2$ -ième colonne pour tout $k \in \{1, \dots, n\}$ contient les $n-k+1$ différences divisées d'ordre k , c'est-à-dire $(f[x_i, x_{i+1}, x_{i+2}, \dots, x_{i+k}])_{0 \leq i \leq n-k}$ données par

$$\forall k \in \{1, \dots, n\}, \quad \forall i \in \{0, \dots, n-k\}, \quad f[x_i, \dots, x_{i+k}] = \frac{f[x_{i+1}, \dots, x_{i+k}] - f[x_i, \dots, x_{i+k-1}]}{x_{i+k} - x_i}. \quad (\text{A.10})$$

- L'avant-dernière colonne contient les 2 différences divisées d'ordre $n-1$, c'est-à-dire, $f[x_0, x_1, \dots, x_{n-1}]$ et $f[x_1, x_2, \dots, x_n]$ définie par (A.10) pour $k = n-1$ et $i \in \{0, 1\}$, c'est-à-dire

$$f[x_0, x_1, \dots, x_{n-1}] = \frac{f[x_1, \dots, x_{n-1}] - f[x_0, x_1, \dots, x_{n-2}]}{x_{n-1} - x_0},$$

$$f[x_1, x_2, \dots, x_n] = \frac{f[x_2, \dots, x_n] - f[x_1, x_2, \dots, x_{n-1}]}{x_n - x_1}.$$

- La dernière colonne contient l'unique différence divisée d'ordre n , c'est-à-dire, $f[x_0, x_1, \dots, x_{n-1}, x_n]$ définie (A.10) pour $k = n$ et $i = 0$, c'est-à-dire

$$f[x_0, x_1, \dots, x_{n-1}, x_n] = \frac{f[x_1, \dots, x_{n-1}, x_n] - f[x_0, x_1, \dots, x_{n-2}, x_{n-1}]}{x_n - x_0}.$$

Voir les tableaux A.1 page ci-contre et A.2 page 66. La double flèche $\begin{matrix} \searrow \\ \nearrow \end{matrix}$ signifie que chaque différence divisée $f[x_i, x_{i+1}, x_{i+2}, \dots, x_{i+k}]$ pour $k \in \{1, \dots, n\}$ et pour $i \in \{0, \dots, n-k\}$ est déterminée de la façon suivante

$$\begin{matrix} A \searrow \\ B \nearrow \end{matrix} \frac{B-A}{x_{i+k} - x_i} = f[x_i, x_{i+1}, x_{i+2}, \dots, x_{i+k}].$$

Enfin, on se sert de ce tableau et uniquement des différences divisées encadrées dans les tableaux A.1 et A.2 pour calculer Π_n donné par

$$\begin{aligned} \Pi_n(x) = & f[x_0] + f[x_0, x_1](x-x_0) + f[x_0, x_1, x_2](x-x_0)(x-x_1) + \dots + f[x_0, x_1, x_2, \dots, x_n](x-x_0)(x-x_1)\dots(x-x_{n-1}), \end{aligned} \quad (\text{A.11a})$$

soit encore

$$\Pi_n(x) = \sum_{i=0}^n f[x_0, \dots, x_i] \prod_{j=0}^{i-1} (x-x_j). \quad (\text{A.11b})$$

avec la convention

$$\prod_{j=0}^{-1} (x-x_j) = 1. \quad (\text{A.12})$$

TABLE A.1. Construction des différences divisées (premières colonnes).

x_i	$f[x_i]$	$f[x_i, x_{i+1}]$	$f[x_i, x_{i+1}, x_{i+2}]$	$f[x_i, x_{i+1}, x_{i+2}, x_{i+3}]$
x_0	$f[x_0] = f(x_0)$			
x_1	$f[x_1] = f(x_1)$	$f[x_0, x_1] = \frac{f[x_1] - f[x_0]}{x_1 - x_0}$	$f[x_0, x_1, x_2] = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0}$	
x_2	$f[x_2] = f(x_2)$	$f[x_1, x_2] = \frac{f[x_2] - f[x_1]}{x_2 - x_1}$	$f[x_1, x_2, x_3] = \frac{f[x_2, x_3] - f[x_1, x_2]}{x_3 - x_1}$	$f[x_0, x_1, x_2, x_3] = \frac{f[x_1, x_2, x_3] - f[x_0, x_1, x_2]}{x_3 - x_0}$
x_3	$f[x_3] = f(x_3)$	$f[x_2, x_3] = \frac{f[x_3] - f[x_2]}{x_3 - x_2}$	$f[x_2, x_3, x_4] = \frac{f[x_3, x_4] - f[x_2, x_3]}{x_4 - x_2}$	$f[x_1, x_2, x_3, x_4] = \frac{f[x_2, x_3, x_4] - f[x_1, x_2, x_3]}{x_4 - x_1}$
x_4	$f[x_4] = f(x_4)$	$f[x_3, x_4] = \frac{f[x_4] - f[x_3]}{x_4 - x_3}$	\vdots	\vdots
\vdots	\vdots	\vdots	\vdots	\vdots
x_{n-1}	$f[x_{n-1}] = f(x_{n-1})$		$f[x_{n-2}, x_{n-1}, x_n] = \frac{f[x_{n-1}, x_n] - f[x_{n-2}, x_{n-1}]}{x_n - x_{n-2}}$	
x_n	$f[x_n] = f(x_n)$	$f[x_{n-1}, x_n] = \frac{f[x_n] - f[x_{n-1}]}{x_n - x_{n-1}}$		

$f[x_i, x_{i+1}, \dots, x_{i+n-2}]$	$f[x_i, x_{i+1}, \dots, x_{i+n-1}]$	$f[x_0, x_1, x_2, \dots, x_n]$
$f[x_0, \dots, x_{n-2}] = \frac{f[x_1, \dots, x_{n-2}] - f[x_0, \dots, x_{n-3}]}{x_{n-2} - x_0}$	$f[x_0, \dots, x_{n-1}] = \frac{f[x_1, \dots, x_{n-1}] - f[x_0, \dots, x_{n-2}]}{x_{n-1} - x_0}$	$f[x_0, \dots, x_n] = \frac{f[x_1, \dots, x_n] - f[x_0, \dots, x_{n-1}]}{x_n - x_0}$
$f[x_1, \dots, x_{n-1}] = \frac{f[x_2, \dots, x_{n-1}] - f[x_1, \dots, x_{n-2}]}{x_{n-1} - x_0}$	$f[x_1, \dots, x_n] = \frac{f[x_2, \dots, x_n] - f[x_1, \dots, x_{n-1}]}{x_n - x_1}$	
$f[x_2, \dots, x_n] = \frac{f[x_3, \dots, x_n] - f[x_2, \dots, x_{n-1}]}{x_n - x_2}$		

TABLE A.2. Construction des différences divisées (dernières colonnes).

THÉORÈME A.4 (Erreur d'interpolation pour des nœuds quelconques). Soient $(x_i)_{0 \leq i \leq n}$ $n + 1$ nœuds distincts dans $[a, b]$ et soit $f \in \mathcal{C}^{n+1}([a, b])$. Alors

$$\forall x \in [a, b], \quad \exists \xi \in]a, b[, \quad f(x) - \Pi_n f(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega_{n+1}(x), \quad (\text{A.13})$$

où ω_{n+1} est défini par

$$\omega_{n+1}(x) = \prod_{i=0}^n (x - x_i). \quad (\text{A.14})$$

THÉORÈME A.5 (Erreur d'interpolation pour des nœuds quelconques). Sous les hypothèses du théorème A.4, on note

$$M_{n+1} = \max_{a \leq x \leq b} |f^{(n+1)}(x)|. \quad (\text{A.15})$$

Alors

$$\forall x \in [a, b], \quad |f(x) - \Pi_n f(x)| \leq \frac{M_{n+1}}{(n+1)!} \max_{x \in [a, b]} |\omega_{n+1}(x)|. \quad (\text{A.16})$$

DÉFINITION A.6 (Nœuds équirépartis). Pour $a < b$ et $n \in \mathbb{N}^*$, $n + 1$ nœuds équirépartis $(x_i)_{0 \leq i \leq n}$ définissent n sous-intervalles de $[a, b]$ de la même taille et sont définis par

$$h = \frac{b - a}{n}, \quad (\text{A.17a})$$

$$\forall i \in \{0, \dots, n\}, \quad x_i = a + ih. \quad (\text{A.17b})$$

THÉORÈME A.7 (Erreur d'interpolation pour des nœuds équirépartis). Pour des nœuds équirépartis, c'est-à-dire donnés par la définition A.6, on a, sous les hypothèses du théorème A.4,

$$E_n(f) = \max_{x \in [a, b]} |f(x) - \Pi_n f(x)| \leq \frac{1}{4(n+1)} \left(\frac{b-a}{n} \right)^{n+1} \max_{x \in [a, b]} |f^{(n+1)}(x)|. \quad (\text{A.18})$$

DÉFINITION A.8. Soit $n \in \mathbb{N}^*$. Sur l'intervalle $[a, b]$, les $n + 1$ points $x_{i0 \leq i \leq n}$ de Tchebycheff sont donnés par

$$\forall i \in \{0, \dots, n\}, \quad z_i = -\cos\left(\frac{\pi i}{n}\right), \quad (\text{A.19a})$$

et

$$\forall i \in \{0, \dots, n\}, \quad x_i = \frac{a+b}{2} + \frac{b-a}{2} z_i, \quad (\text{A.19b})$$

REMARQUE A.9. Notons aussi l'interprétation géométrique des formules (A.19) : on considère le demi-cercle de centre d'abscisse $(a+b)/2$ et de rayon $(b-a)/2$, c'est-à-dire passant par les points d'abscisses a et b . On divise ce demi-cercle en n parties égales ; on a donc $n + 1$ points définis par les angles $i\pi/n$ pour $0 \leq i \leq n$. Voir la figure A.1 page suivante. Les $n + 1$ abscisses de ces points ne sont autres que les points définis par (A.19).

REMARQUE A.10. Attention, parfois une définition légèrement différente de (A.19a) est donnée (c'est le cas dans [BM03, p. 282]) :

$$\forall i \in \{0, \dots, n\}, \quad z_i = \cos\left(\frac{2i+1}{n+1} \frac{\pi}{2}\right). \quad (\text{A.20})$$

◇

THÉORÈME A.11 (Erreur d'interpolation composite). Soient $x_0 = A < x_1 < \dots < x_N = B$ des points qui divisent $I = [A, B]$. On note $I_j = [x_{j-1}, x_j]$ les sous-intervalles de longueur h_j et $h = \max_{1 \leq j \leq N} h_j$. Sur chaque

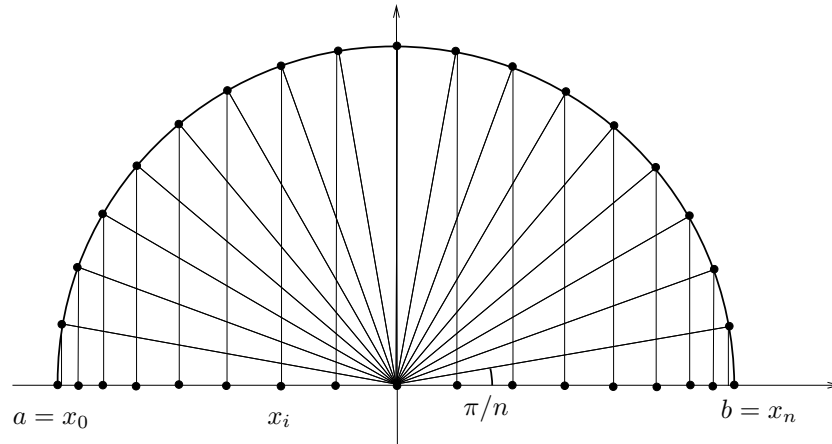


FIGURE A.1. Les points de Tchebycheff.

sous-intervalle I_j , on interpole $f|_{I_j}$ par un polynôme de degré² n avec des points équirépartis. Le polynôme par morceaux est noté $\Pi_n^h f(x)$. Si $f \in \mathcal{C}^{n+1}([A, B])$, alors, on a

$$E_n^h(f) = \max_{x \in [A, B]} |f(x) - \Pi_n^h f(x)| \leq \frac{1}{4(n+1)n^{n+1}} \max_{x \in [A, B]} |f^{(n+1)}(x)| h^{n+1}. \quad (\text{A.21})$$

Si de plus, les points sont équirépartis, on a

$$E_n^h(f) = \max_{x \in [A, B]} |f(x) - \Pi_n^h f(x)| \leq \frac{(B-A)^{n+1}}{4(n+1)n^{n+1}} \max_{x \in [A, B]} |f^{(n+1)}(x)| \frac{1}{N^{n+1}}. \quad (\text{A.22})$$

A.2. Intégration

méthode	formule
rectangle (gauche)	$(b-a)f(a)$
rectangle (droite)	$(b-a)f(b)$
milieu	$(b-a)f((a+b)/2)$
trapèze	$\frac{1}{2}(b-a)(f(a) + f(b))$
Simpson	$\frac{1}{6}(b-a)(f(a) + 4f((a+b)/2) + f(b))$

TABLE A.3. Méthodes élémentaires sur $[a, b]$.

A.3. Équations non-linéaires

LEMME A.12. Soit f continue sur un intervalle $[a, b]$ telle que

$$f(a)f(b) \leq 0. \quad (\text{A.23})$$

Alors, il existe un zéro de f sur $[a, b]$.

2. Attention, N le nombre de sous-intervalles ou morceaux n'a aucun lien avec n le degré des polynômes !

méthode	formule
rectangle (gauche)	$h \sum_{i=0}^{N-1} f(x_i)$
rectangle (droite)	$h \sum_{i=1}^N f(x_i)$
milieu	$h \sum_{i=0}^{N-1} f(x_i + h/2)$
trapèze	$\frac{h}{2}(f(A) + f(B)) + h \sum_{i=1}^{N-1} f(x_i)$
Simpson	$\frac{h}{6} \left(f(A) + f(B) + 2 \sum_{i=1}^{N-1} f(x_i) + 4 \sum_{i=0}^{N-1} f\left(x_i + \frac{h}{2}\right) \right)$

TABLE A.4. Méthodes composites (composées) sur $[A, B]$. Dans ce tableau, $N \in \mathbb{N}^*$, $h = (B - A)/N$ et, pour tout $i \in \{0, \dots, N\}$, $x_i = a + ih$.

méthode	erreur
rectangle	$\frac{(b-a)^2}{2} f'(\eta)$
milieu	$\frac{(b-a)^3}{24} f''(\eta)$
trapèze	$-\frac{(b-a)^3}{12} f''(\eta)$
Simpson	$-\frac{(b-a)^5}{2880} f^{(4)}(\eta)$

TABLE A.5. Erreurs des méthodes élémentaires sur $[a, b]$. Dans ce tableau, η appartient à $]a, b[$.

méthode	erreur
rectangle	$h \frac{B-A}{2} f'(\eta)$
milieu	$h^2 \frac{B-A}{24} f''(\eta)$
trapèze	$-h^2 \frac{B-A}{12} f''(\eta)$
Simpson	$-h^4 \frac{B-A}{2880} f^{(4)}(\eta)$

TABLE A.6. Erreurs des méthodes composites (composées) sur $[A, B]$. Dans ce tableau, η appartient à $[A, B]$.

LEMME A.13. Soit f vérifiant les hypothèses du lemme A.12. La méthode de dichotomie est convergente. Autrement dit, la suite (x_n) construite converge vers un des zéros de f .

PROPOSITION A.14. Avec l'algorithme de dichotomie (sous les hypothèses du lemme A.12), si on cherche la racine r avec une erreur absolue ε (précision recherchée), il faut effectuer n itérations telles que

$$\frac{b-a}{2^n} \leq \varepsilon, \tag{A.24}$$

soit encore

$$n \geq \frac{\ln((b-a)/\varepsilon)}{\ln 2}, \quad (\text{A.25})$$

soit encore,

$$n = \left\lceil \frac{\ln((b-a)/\varepsilon)}{\ln 2} \right\rceil, \quad (\text{A.26})$$

où³, pour tout réel X ,

$$\lceil X \rceil \text{ est le plus petit entier supérieur ou égal à } X. \quad (\text{A.27})$$

DÉFINITION A.15 (Méthode du point fixe). La suite $(x_n)_{n \in \mathbb{N}}$ est définie par la donnée de $x_0 \in \mathbb{R}$ et

$$\forall n \in \mathbb{N}, \quad x_{n+1} = g(x_n). \quad (\text{A.28})$$

PROPOSITION A.16 (condition suffisante de divergence de la méthode du point fixe). *Supposons que g soit dérivable en r , qu'*

$$\text{il existe une infinité d'indices } n \text{ tels que } x_n \neq r, \quad (\text{A.29a})$$

et que

$$|g'(r)| > 1. \quad (\text{A.29b})$$

Alors, il existe un intervalle du type $I = [r - \varepsilon_0, r + \varepsilon_0]$, tel que, pour tout x_0 dans $I \setminus \{r\}$, la méthode du point fixe (définition A.15) ne converge pas.

PROPOSITION A.17 (condition suffisante de divergence de la méthode du point fixe). *Supposons qu'il existe un intervalle $I = [a, b]$ de \mathbb{R} tel que $r \in I$ et sur lequel g vérifie les trois hypothèses suivantes :*

$$g \text{ est définie sur } I, \quad (\text{A.30a})$$

$$\text{pour tout } x \notin I, \text{ si } g(x) \text{ est défini, il n'appartient pas à } I, \quad (\text{A.30b})$$

$$\text{il existe un réel } k > 1 \text{ tel que } \forall x \in I, \quad |g'(x)| \geq k. \quad (\text{A.30c})$$

Alors, pour tout x_0 dans $I \setminus \{r\}$, la méthode du point fixe ne converge pas.

DÉFINITION A.18 (Intervalle g -stable). Si g est une fonction définie au moins sur un intervalle I de \mathbb{R} , on dit que I est g -stable ssi $g(I) \subset I$, ce qui est aussi équivalent à pour tout x dans I , $g(x)$ appartient à I .

PROPOSITION A.19 (condition suffisante de convergence de la méthode du point fixe). *Supposons qu'il existe un intervalle $I = [a, b]$ de \mathbb{R} sur lequel g vérifie les deux hypothèses suivantes :*

$$g \text{ est définie sur } I \text{ et } I \text{ est } g\text{-stable}; \quad (\text{A.31a})$$

$$g \text{ est dérivable sur } I \text{ et il existe un réel } k \text{ de } [0, 1[\text{ tel que } \forall x \in I, \quad |g'(x)| \leq k. \quad (\text{A.31b})$$

Alors, on a les deux résultats suivants :

- (1) g admet un point fixe unique r dans $I = [a, b]$;
- (2) pour tout x_0 de I , la suite (x_n) est définie et converge vers r .

PROPOSITION A.20 (Majoration de l'erreur pour la méthode du point fixe). *Sous les hypothèses de la proposition A.19, on a*

$$\forall n \in \mathbb{N}, \quad |x_n - r| \leq k^n (b - a). \quad (\text{A.32})$$

3. En anglais, on parle de "plafond" : ceil.

PROPOSITION A.21. Avec la méthode du point fixe (sous les hypothèses de la proposition A.19), si on cherche le point fixe r avec une erreur absolue ε (précision recherchée), il faut effectuer n itérations telles que

$$k^n (b - a) \leq \varepsilon, \quad (\text{A.33})$$

soit encore

$$n \geq -\frac{\ln((b-a)/\varepsilon)}{\ln k}, \quad (\text{A.34})$$

soit encore,

$$n = \left\lceil -\frac{\ln((b-a)/\varepsilon)}{\ln k} \right\rceil, \quad (\text{A.35})$$

où $\lceil \cdot \rceil$ est défini par (A.27).

DÉFINITION A.22. Soit $p \in \mathbb{N}^*$. Nous dirons qu'une méthode de point fixe (A.28) est d'ordre au moins p si, l'erreur $e_n = x_n - r$ vérifie⁴ :

$$\exists C \in \mathbb{R}_+, \quad \forall n \in \mathbb{N}, \quad \frac{|e_{n+1}|}{|e_n|^p} \leq C. \quad (\text{A.37})$$

Elle est dite exactement d'ordre p si

$$\lim_{n \rightarrow +\infty} \frac{|e_{n+1}|}{|e_n|^p} \neq 0. \quad (\text{A.38})$$

LEMME A.23. Si une méthode de point fixe est au moins d'ordre $p \in \mathbb{N}^*$ (au sens de la définition A.22), alors

- si $p = 1$, on a

$$\forall n \geq 0, \quad |e_n| \leq |e_0|C^n, \quad (\text{A.39})$$

- si $p > 1$, alors en définissant les nombres γ et δ par

$$\gamma = C^{\left(\frac{1}{1-p}\right)}, \quad (\text{A.40a})$$

$$\delta = |e_0|C^{\left(\frac{1}{p-1}\right)}, \quad (\text{A.40b})$$

on a

$$\forall n \geq 0, \quad |e_n| \leq \gamma \delta^{(p^n)}. \quad (\text{A.41})$$

PROPOSITION A.24 (Majoration de l'erreur pour une méthode d'ordre p). Si une méthode de point fixe est au moins d'ordre $p \in \mathbb{N}^*$ et :

- si $p = 1$ et si $C < 1$
- si $p > 1$ et si $|e_0|$ est assez petit de telle sorte que pour que

$$|e_0|C^{\left(\frac{1}{p-1}\right)} < 1, \quad (\text{A.42})$$

la suite x_n tend vers r . Sous ces hypothèses, si on cherche le point fixe r avec une erreur absolue ε (précision recherchée), il faut effectuer n itérations avec n défini par que

- si $p = 1$

$$n = \left\lceil -\frac{\ln(|e_0|/\varepsilon)}{\ln C} \right\rceil, \quad (\text{A.43})$$

- si $p > 1$

$$n = \left\lceil \frac{1}{\ln p} \ln \left(\frac{\ln \frac{\varepsilon}{\gamma}}{\ln \delta} \right) \right\rceil, \quad (\text{A.44})$$

avec γ et δ définis par (A.40).

On rappelle que $\lceil \cdot \rceil$ est défini par (A.27).

◇

4. On pourra remplacer ceci par l'inégalité un peu moins exigeante :

$$\exists C \in \mathbb{R}_+, \quad \exists N \in \mathbb{N}, \quad \forall n \geq N, \quad \frac{|e_{n+1}|}{|e_n|^p} \leq C. \quad (\text{A.36})$$

PROPOSITION A.25. Si g est de classe \mathcal{C}^1 sur un intervalle du type $[r - \varepsilon, r + \varepsilon]$, avec ε assez petit, si x_0 appartient à $[r - \varepsilon, r + \varepsilon]$, si $g(r) = r$ et si

$$|g'(r)| < 1, \quad (\text{A.45})$$

alors la méthode de point fixe est au moins linéaire, c'est-à-dire que la définition A.22 est vérifiée pour $p = 1$. Si, de plus,

$$g'(r) \neq 0, \quad (\text{A.46})$$

alors la méthode de point fixe est exactement linéaire.

PROPOSITION A.26. Si g est de classe \mathcal{C}^2 sur un intervalle du type $[r - \varepsilon, r + \varepsilon]$, avec ε assez petit, si x_0 appartient à $[r - \varepsilon, r + \varepsilon]$, si $g(r) = r$ et si

$$g'(r) = 0. \quad (\text{A.47})$$

alors la méthode de point fixe est au moins quadratique, c'est-à-dire que la définition A.22 est vérifiée pour $p = 2$. Si, de plus,

$$g''(r) \neq 0, \quad (\text{A.48})$$

alors la méthode de point fixe est exactement quadratique.

PROPOSITION A.27. Si g est de classe \mathcal{C}^3 sur un intervalle du type $[r - \varepsilon, r + \varepsilon]$, avec ε assez petit, si x_0 appartient à $[r - \varepsilon, r + \varepsilon]$, si $g(r) = r$ et si

$$g'(r) = 0, \quad (\text{A.49a})$$

$$g''(r) = 0, \quad (\text{A.49b})$$

alors la méthode de point fixe est au moins cubique, c'est-à-dire que la définition A.22 est vérifiée pour $p = 3$. Si, de plus,

$$g'''(r) \neq 0, \quad (\text{A.50})$$

alors la méthode de point fixe est exactement cubique.

PROPOSITION A.28. Soit $p \in \mathbb{N}^*$. Si g est de classe \mathcal{C}^p sur un intervalle du type $[r - \varepsilon, r + \varepsilon]$, avec ε assez petit, si x_0 appartient à $[r - \varepsilon, r + \varepsilon]$, si $g(r) = r$ et si

$$\forall q \in \{1, \dots, p-1\}, \quad g^{(q)}(r) = 0, \quad (\text{A.51})$$

alors la méthode de point fixe est d'ordre au moins p , c'est-à-dire que la définition A.22 est vérifiée pour p . Si, de plus,

$$g^{(p)}(r) \neq 0, \quad (\text{A.52})$$

alors la méthode de point fixe est d'ordre exactement p .

◇

DÉFINITION A.29 (Méthode de Newton). La suite $(x_n)_{n \in \mathbb{N}}$ est définie par la donnée de $x_0 \in \mathbb{R}$ et

$$\forall n \in \mathbb{N}, \quad x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \quad (\text{A.53})$$

PROPOSITION A.30. Si f est de classe \mathcal{C}^3 sur un intervalle du type $[r - \varepsilon, r + \varepsilon]$, si $f(r) = 0$, $f'(r) \neq 0$ et $f''(r) \neq 0$, alors la méthode de Newton est exactement quadratique.

DÉFINITION A.31 (Méthode de la Sécante). La suite $(x_n)_{n \in \mathbb{N}}$ est définie par les données de $x_0, x_1 \in \mathbb{R}$ et

$$\forall n \in \mathbb{N}^*, \quad x_{n+1} = x_n - \frac{f(x_n)(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})}. \quad (\text{A.54})$$

A.4. Équations différentielles

DÉFINITION A.32. Pour tout $T > 0$, pour tout $\xi_0 \in \mathbb{R}$, pour toute fonction f de $[0, T] \times \mathbb{R}$ dans \mathbb{R} , on considère l'équation différentielle ordinaire :

$$\forall t \in [0, T], \quad y'(t) = f(t, y(t)), \quad (\text{A.55a})$$

avec la condition initiale

$$y(0) = \xi_0. \quad (\text{A.55b})$$

DÉFINITION A.33. Pour pour $T > 0$ et pour tout $N \in \mathbb{N}^*$, on pose

$$h = \frac{T}{N}, \quad (\text{A.56a})$$

$$\forall n \in \{0, \dots, N\}, \quad t_n = hn. \quad (\text{A.56b})$$

DÉFINITION A.34 (Schéma numérique d'Euler explicite (ou progressif)). Avec les notations des définitions A.32 et A.33, les approximations du schéma numérique d'Euler explicite (ou progressif), $(y_n)_{0 \leq n \leq N}$ sont données par

$$y_0 = \xi_0, \quad (\text{A.57a})$$

$$\forall n \in \{0, \dots, N-1\}, \quad y_{n+1} = y_n + hf(t_n, y_n). \quad (\text{A.57b})$$

DÉFINITION A.35 (Schéma numérique d'Euler implicite (ou rétrograde)). Avec les notations des définitions A.32 et A.33, les approximations du schéma numérique d'Euler implicite (ou rétrograde), $(y_n)_{0 \leq n \leq N}$ sont données par

$$y_0 = \xi_0, \quad (\text{A.58a})$$

et, pour tout $n \in \{0, \dots, N-1\}$, y_{n+1} vérifie l'équation (non linéaire)

$$y_{n+1} = y_n + hf(t_{n+1}, y_{n+1}). \quad (\text{A.58b})$$

DÉFINITION A.36 (Schéma numérique de Runge-Kutta 2). Avec les notations des définitions A.32 et A.33, les approximations du schéma numérique de Runge-Kutta 2, $(y_n)_{0 \leq n \leq N}$ sont données par

$$y_0 = \xi_0, \quad (\text{A.59a})$$

$$\forall n \in \{0, \dots, N-1\}, \quad \begin{cases} k_n^{(1)} = hf(t_n, y_n), \\ k_n^{(2)} = hf(t_n + h, y_n + k_n^{(1)}), \\ y_{n+1} = y_n + \frac{1}{2}(k_n^{(1)} + k_n^{(2)}). \end{cases} \quad (\text{A.59b})$$

DÉFINITION A.37 (Schéma numérique de Runge-Kutta 4). Avec les notations des définitions A.32 et A.33, les approximations du schéma numérique de Runge-Kutta 4, $(y_n)_{0 \leq n \leq N}$ sont données par

$$y_0 = \xi_0, \quad (\text{A.60a})$$

$$\forall n \in \{0, \dots, N-1\}, \quad \begin{cases} k_n^{(1)} = hf(t_n, y_n), \\ k_n^{(2)} = hf\left(t_n + \frac{1}{2}h, y_n + \frac{1}{2}k_n^{(1)}\right), \\ k_n^{(3)} = hf\left(t_n + \frac{1}{2}h, y_n + \frac{1}{2}k_n^{(2)}\right), \\ k_n^{(4)} = hf\left(t_n + h, y_n + k_n^{(3)}\right), \\ y_{n+1} = y_n + \frac{1}{6}(k_n^{(1)} + 2k_n^{(2)} + 2k_n^{(3)} + k_n^{(4)}). \end{cases} \quad (\text{A.60b})$$

Simulations numériques sur l'erreur d'interpolation

Pour f et g données par

$$\forall x \in [0, 1], \quad f(x) = \sin(x/3), \quad (\text{B.1a})$$

$$\forall x \in [6000, 6001], \quad g(x) = \sin((x - 6000)/3), \quad (\text{B.1b})$$

on calcule le polynôme p_n correspondant aux points équirépartis x_i de $[a, b]$ (c'est-à-dire vérifiant $p_n(x_i) = f(x_i)$) par l'une des trois méthodes données en page 9.

Voir la figure B.1 pour laquelle les trois méthodes de calculs se déroulent bien : f et p_n sont indiscernables graphiquement.

Pour n plus grand, la méthode utilisant la forme canonique n'est plus efficace, alors que les deux autres le sont encore ! Voir la figure B.2.

Pour n encore plus grand, la méthode utilisant les polynômes de Lagrange commence à n'être plus efficace, alors que la méthode avec Newton avec évaluation d'Horner l'est encore. Ensuite, au delà d'une certaine valeur de n cette méthode-ci ne marche plus ! Voir la figure B.3.

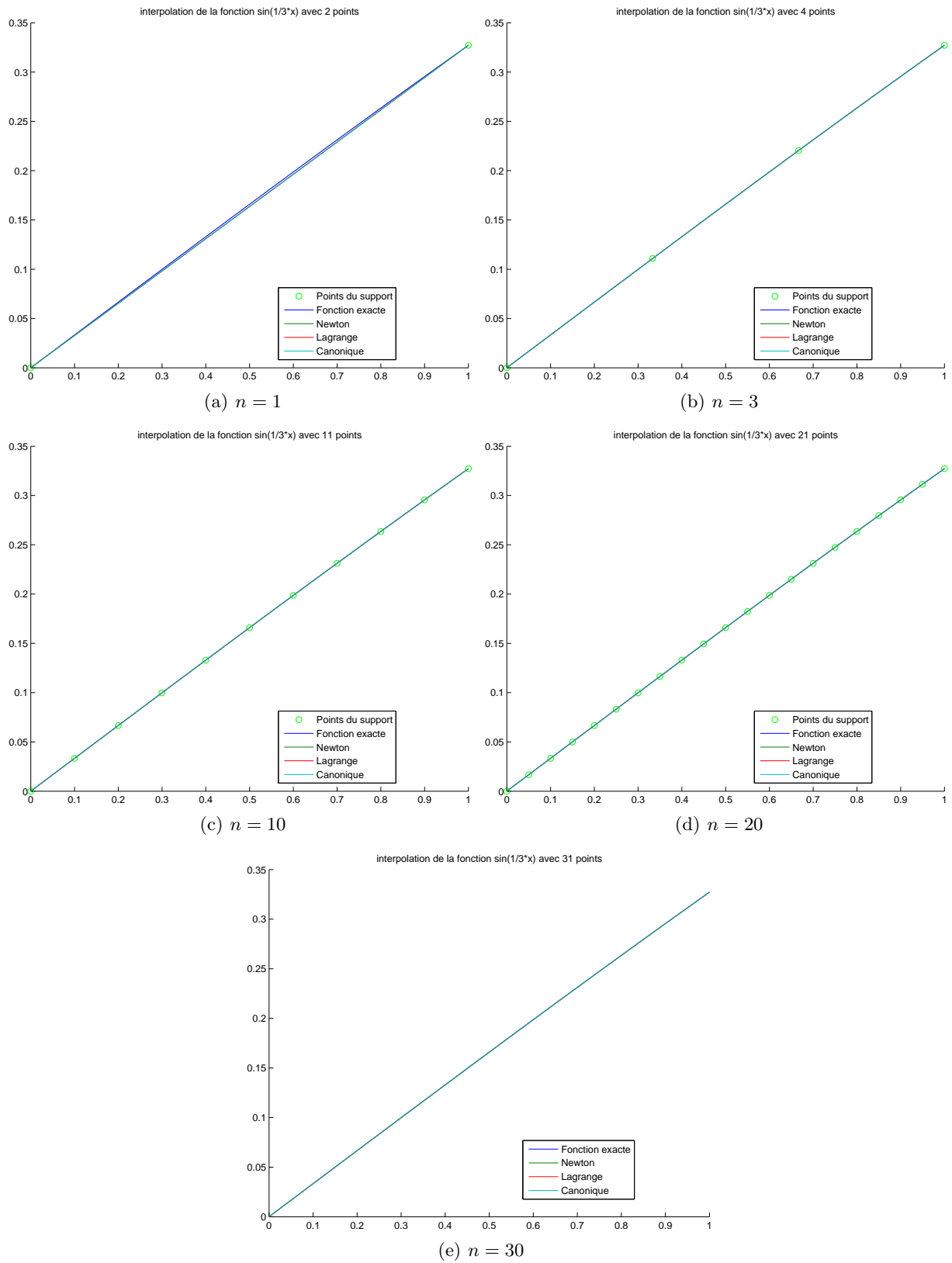
Enfin, on peut aussi tracer le logarithme en base 10 de l'erreur entre f et son interpolé en fonction de n comme le montre la figure B.4. On constate que celle-ci est minimale autour de la valeur $n_0 = 9$ et que les trois méthodes coïncident rigoureusement jusqu'à n_0 . Ensuite, de $n_0 + 1$ à $n_1 = 25$, les deux méthodes Newton et canonique voient leur erreur légèrement remonter, en contradiction avec la théorie, à cause des arrondis de calculs. Il en est de même pour celle correspondant à la base Lagrange, avec une erreur un peu moins petite. Ensuite, au-delà de n_1 , l'erreur en canonique augmente plus rapidement que les deux autres méthodes, qui se comportent mieux. Notons aussi la légère supériorité en terme d'erreur de la forme de Newton, qui sera donc à privilégier tout le temps, compte tenu de sa simplicité algorithmique.

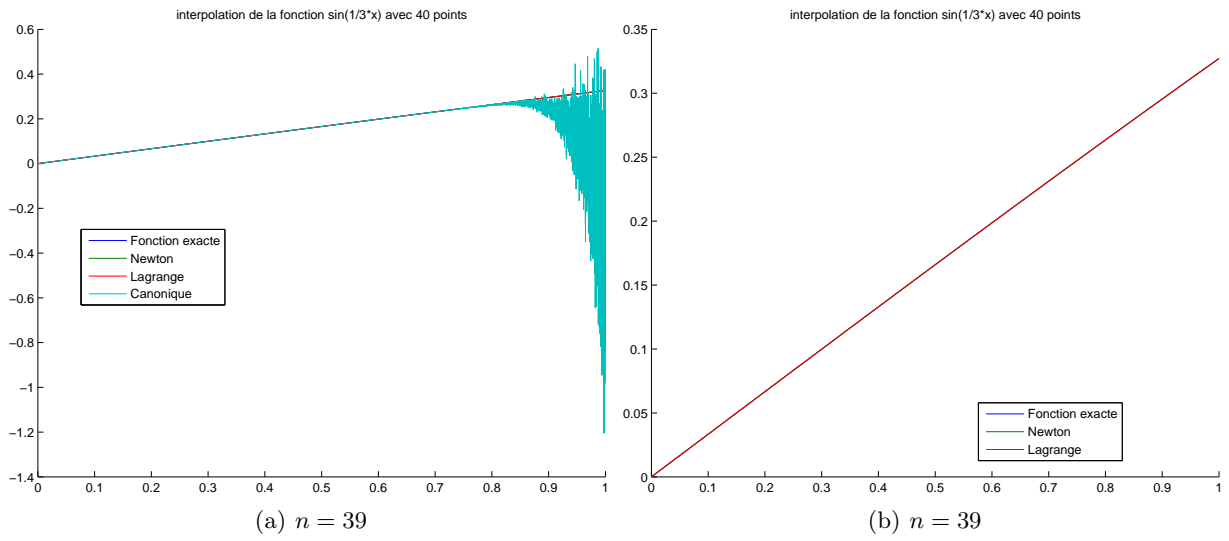
Dans un second temps, on peut aussi refaire les mêmes calculs portant cette fois-ci sur la fonction g définie par (B.1b) on obtient un comportement similaire de l'erreur.

Ici, on constate que la méthode de la forme canonique se comporte beaucoup plus mal que dans le cas précédent. Voir la figure B.5. Dès la valeur de $n_0 = 5$, elle commence à ne plus fonctionner, alors que les deux autres se comportent bien ! Ensuite, les méthodes fondées sur Newton et Lagrange donnent encore de bons résultats, du même ordre, Newton se comportant un peu mieux que Lagrange.

On peut aussi tracer le logarithme en base 10 de l'erreur entre g et son interpolé en fonction de n comme le montre la figure B.6. Le comportement de l'erreur est identique à celui de la figure B.4, mis à part le fait que la méthode Lagrange a une erreur numérique qui apparaît pour des valeurs de n beaucoup plus petite que précédemment.

Finalement, retenons que la forme de Newton et l'évaluation par l'algorithme de Horner évoqués dans le point 1 page 9 est largement supérieure aux deux autres méthodes (Lagrange et base canonique) tant sur le plan algorithmique que numérique.

FIGURE B.1. La fonction f et son polynôme interpolateur p_n calculés de trois façons différentes.

FIGURE B.2. La fonction f et son polynôme interpolateur p_n calculés de deux façons différentes.

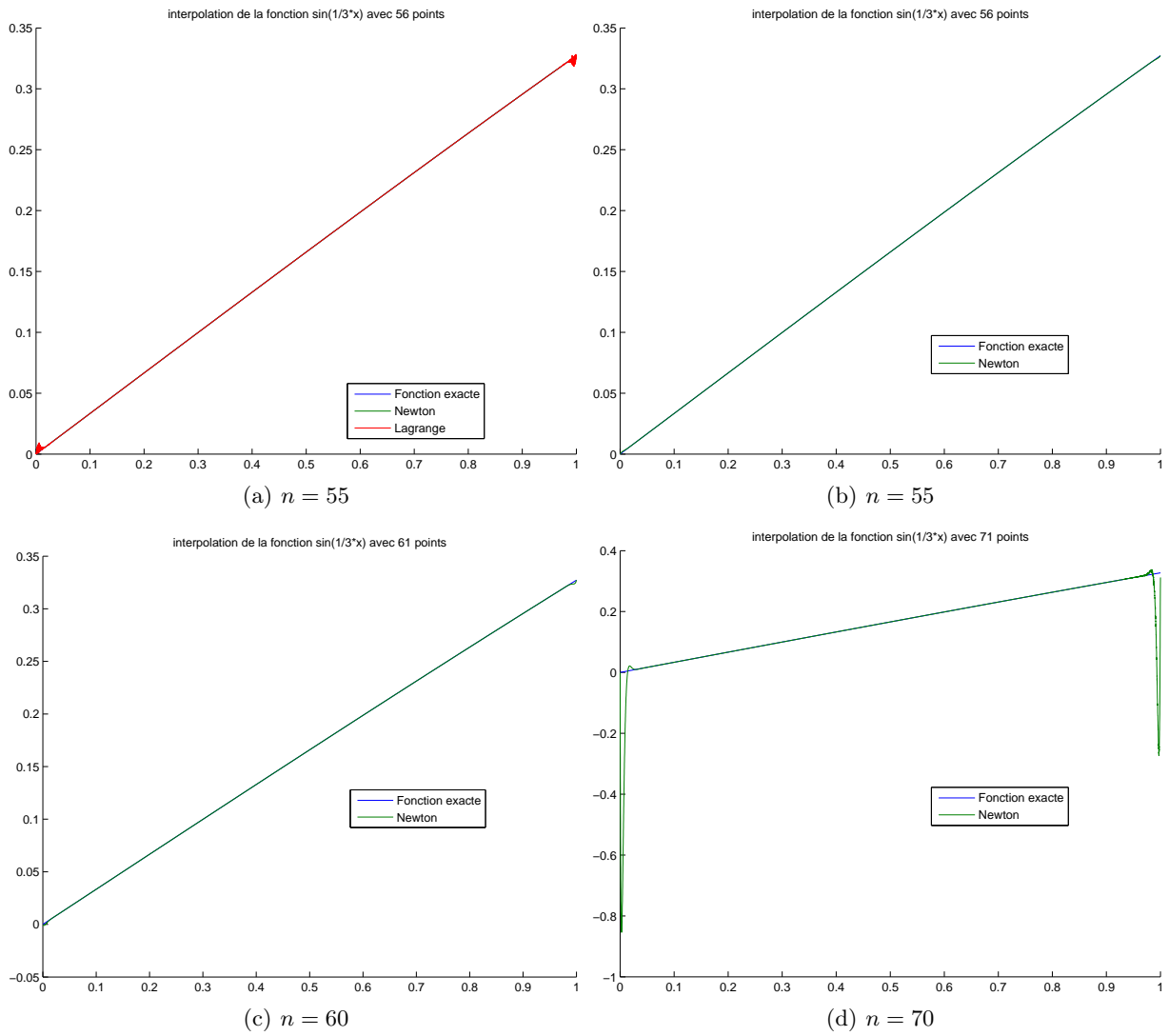
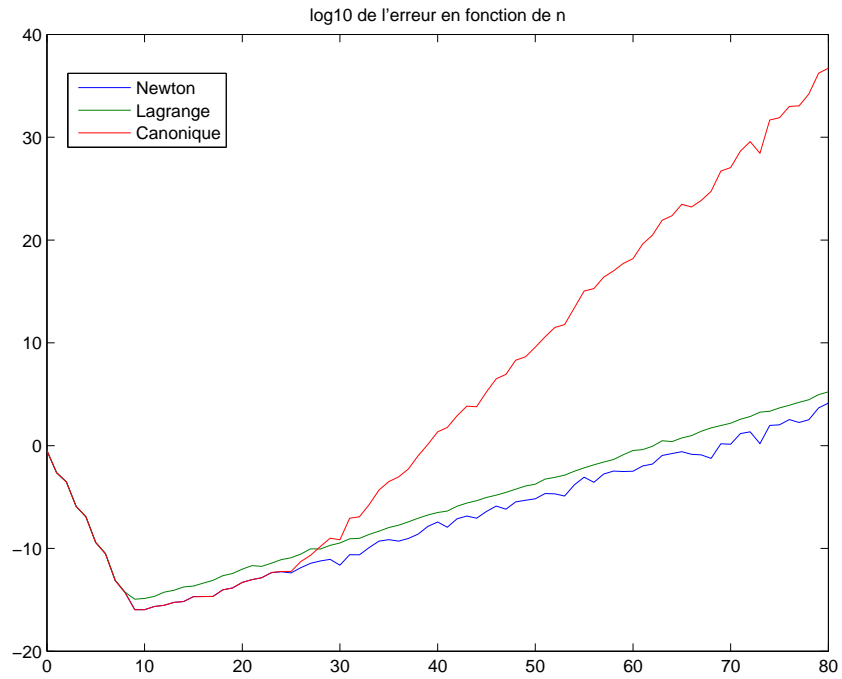


FIGURE B.3. La fonction f et son polynôme interpolateur p_n calculés de une ou deux façons différentes.

FIGURE B.4. Logarithme de l'erreur en base 10 entre f et son interpolée p_n .

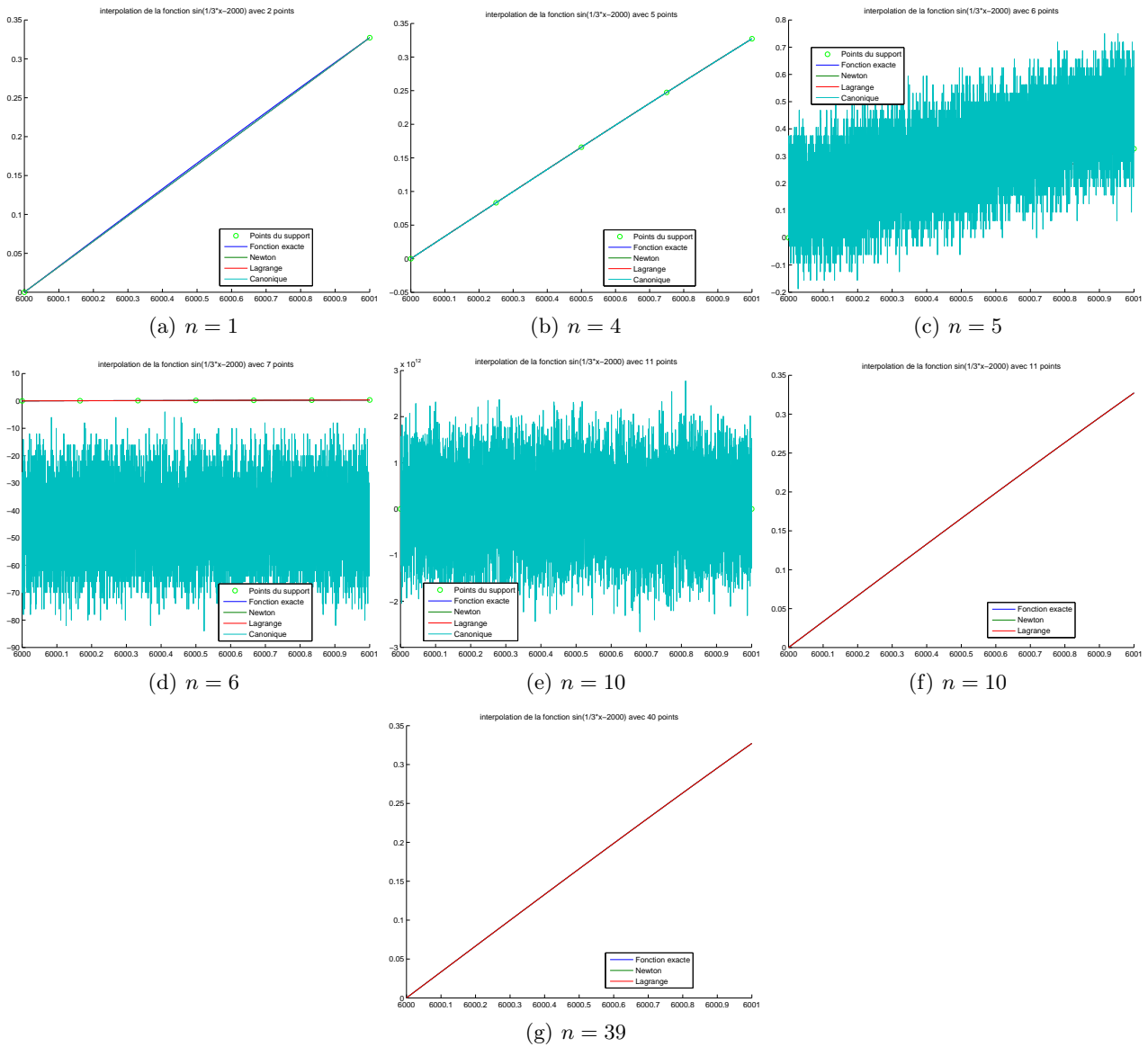
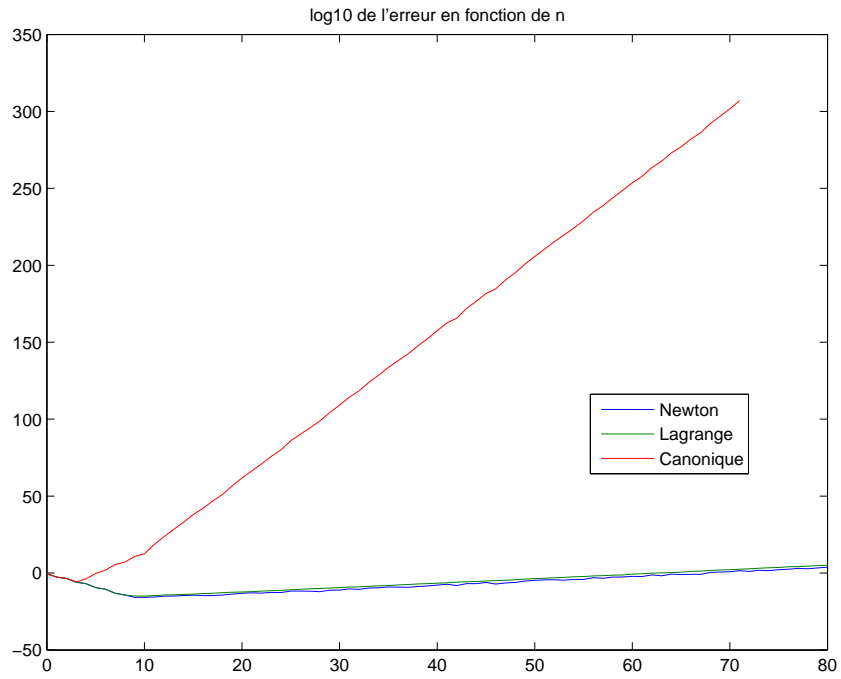


FIGURE B.5. La fonction g et son polynôme interpolateur p_n calculés de trois façons différentes.

FIGURE B.6. Logarithme de l'erreur en base 10 entre g et son interpolée p_n .

Étude du majorant exact d'une dérivée quatrième

On pourra consulter et faire tourner la fonction matlab `corrige_majoration_fderivee4.m` disponible à l'url habituelle, qui fournit successivement tous les résultats de cette annexe.

On cherche à étudier les extremas de la dérivée quatrième de la fonction définie par

$$\forall x \in [0, 1], \quad f(x) = e^{-x^2}, \quad (\text{C.1})$$

donnée par (2.10). Il faut donc étudier la fonction donnée par (après division par 4) :

$$\forall x \in [0, 1], \quad g(x) = e^{-x^2} (3 - 12x^2 + 4x^4). \quad (\text{C.2})$$

(1) Si on utilise la fonction fournie sur le site habituel `maxabsfun`, on tape

```
[Ms,M,Md]=maxabsfun(0,1,@(x)exp(-x.^2).*(3-12.*x.^2+4.*x.^4),[],1);
```

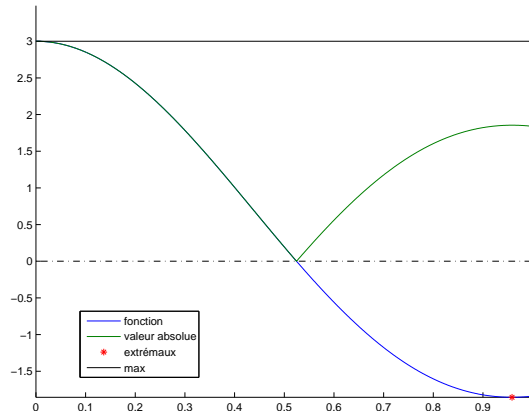


FIGURE C.1. Les fonctions g et $|g|$ pour g définie par (C.2).

On obtient la figure C.1, sur laquelle on constate que le maximum de la valeur absolue de g est donné par

$$M = 3. \quad (\text{C.3})$$

Cette preuve n'en est pas une car cette observation n'est pas justifiée pour l'instant.

(2) Essayons de procéder comme dans la remarque 1.3 page 11 de l'exercice 1.4. De façon analogue à l'inégalité (1.31), on écrit grâce à (C.2) :

$$\forall x \in [0, 1], \quad g(x) = e^{-x^2} (3 - 12x^2 + 4x^4),$$

dont on déduit comme dans (1.32) et (1.33) :

$$\forall x \in [0, 1], \quad |g(x)| = \left| e^{-x^2} \right| |p(x)|,$$

avec

$$p(x) = 3 - 12x^2 + 4x^4, \quad (\text{C.4})$$

et donc

$$\forall x \in [0, 1], \quad |g(x)| \leq |p(x)|.$$

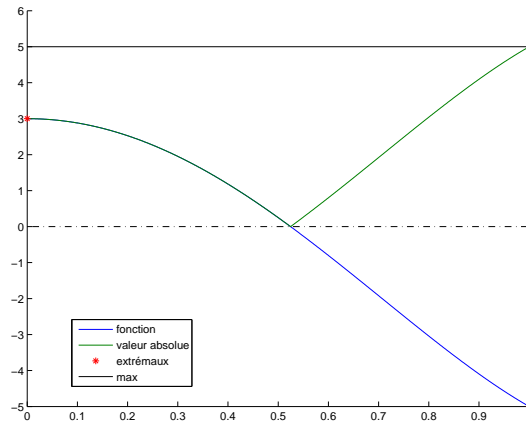


FIGURE C.2. Les fonctions p et $|p|$ pour p définie par (C.4).

En utilisant la fonction `maxabsfun`, on obtient la figure C.2, sur laquelle on constate que le maximum de la valeur absolue de p vaut 5, ce qui donnerait une majoration de $|g|$ donnée par

$$M = 20,$$

ce qui est trop grand par rapport à (C.3).

- (3) Étudions la dérivée de g de façon plus rigoureuse que dans la question 1. Il n'est pas nécessaire d'étudier totalement la fonction g pour en donner les extremas de la valeur absolue. Il suffit d'observer qu'une fonction dérivable sur un segment $[a, b]$ atteint ses bornes, soit aux bords a et b , soit en des points intérieurs à $]a, b[$ où la dérivée de cette fonction s'annule. Ainsi, la valeur absolue maximale de cette fonction est la plus grande valeur de $|g|$ aux points a , b et aux zéros intérieurs de la dérivée de cette fonction. Ce sont en fait, exactement les calculs que mènent la fonction `maxabsfun.m`. Dans la mesure où les résultats sont ici obtenus en symbolique, c'est-à-dire de façon exacte, on peut donc affirmer *a posteriori* que les calculs de la question 1 sont exactement prouvés !

La dérivée de g vaut :

$$g'(x) = 2xe^{-x^2} (-15 + 20x^2 - 4x^4), \quad (\text{C.5})$$

dont les cinq racines sont données brutalement par Matlab :

$$x_0 = 0; \quad (\text{C.6a})$$

$$x_1 = -1/2 \sqrt{10 - 2\sqrt{10}}; \quad (\text{C.6b})$$

$$x_2 = 1/2 \sqrt{10 - 2\sqrt{10}}; \quad (\text{C.6c})$$

$$x_3 = -1/2 \sqrt{10 + 2\sqrt{10}}; \quad (\text{C.6d})$$

$$x_4 = 1/2 \sqrt{10 + 2\sqrt{10}}; \quad (\text{C.6e})$$

et dont les valeurs sont

$$x_0 = 0; \quad (\text{C.7a})$$

$$x_1 = -0.9585724646138; \quad (\text{C.7b})$$

$$x_2 = 0.9585724646138; \quad (\text{C.7c})$$

$$x_3 = -2.0201828704561; \quad (\text{C.7d})$$

$$x_4 = 2.0201828704561. \quad (\text{C.7e})$$

La seule racine appartenant à l'intervalle $]0, 1[$ est x_2 . Les valeurs prises par g en ce point ces points et les bords 0 et 1 sont données par

$$\left\{ 3, -5e^{-1}, -4e^{-5/2+1/2\sqrt{10}}(-2+\sqrt{10}) \right\},$$

correspondant à

$$\{3.00000000000000, -1.83939720585721, -1.85487029414460\}.$$

La plus grande en valeur absolue est 3, ce qui prouve finalement (C.3).

- (4) On peut préciser cela en remarquant que le polynôme qui intervient dans l'expression (C.5) de la dérivée de g est un polynôme bicarré correspondant à

$$q(x) = -15 + 20x - 4x^2,$$

dont les *deux* racines sont données par :

$$y_0 = 5/2 - 1/2\sqrt{10}; \quad (\text{C.8a})$$

$$y_1 = 5/2 + 1/2\sqrt{10}, \quad (\text{C.8b})$$

et dont les valeurs sont

$$y_0 = 0.9188611699158; \quad (\text{C.9a})$$

$$y_1 = 4.0811388300842. \quad (\text{C.9b})$$

La seule racine appartenant à l'intervalle $]0, 1[$ est y_0 . En reprenant la racine carrée de ce nombre, on retrouve (C.6) et (C.7) et on conclue comme précédemment.

Définition et utilisation de la fonction W de Lambert

On pourra consulter par exemple https://fr.wikipedia.org/wiki/Fonction_W_de_Lambert

D.1. Définition de la fonction W de Lambert

On cherche à résoudre l'équation suivante : pour $z \in \mathbb{R}$ donné, on cherche $w \in \mathbb{R}$ tel que

$$we^w = z. \quad (\text{D.1})$$

On a alors le résultat suivant :

PROPOSITION D.1 (les deux branches W_0 et W_{-1} de la "fonction" de Lambert).

- Si $z < -\frac{1}{e}$, il n'existe aucun $w \in \mathbb{R}$ solution de (D.1) ;
- Il existe une unique fonction W_0 de $[-\frac{1}{e}, +\infty[$ dans $[-1, +\infty[$ telle que (D.1) est équivalent à $w = W_0(z)$;
- Il existe une unique fonction W_{-1} de $[-\frac{1}{e}, 0[$ dans $]-\infty, -1]$ telle que (D.1) est équivalent à $w = W_{-1}(z)$.

Cela est équivalent à dire : Pour tout $z \in \mathbb{R}$:

- Si $z < -\frac{1}{e}$, l'équation (D.1) n'a pas de solution ;
- Si $z = -\frac{1}{e}$, la seule solution de (D.1) est donnée par

$$w = W_{-1}\left(-\frac{1}{e}\right) = W_0\left(-\frac{1}{e}\right) = -1; \quad (\text{D.2})$$

- Si $z \in]-\frac{1}{e}, 0[$, les deux solutions distinctes de (D.1) sont données par

$$w = W_{-1}(z), \quad (\text{D.3a})$$

$$w = W_0(z); \quad (\text{D.3b})$$

- Si $z \in [0, +\infty[$, l'unique solution de (D.1) est donnée par

$$w = W_0(z). \quad (\text{D.4})$$

DÉMONSTRATION. Il suffit d'étudier la fonction f définie sur \mathbb{R} par

$$\forall w \in \mathbb{R}, \quad f(w) = we^w, \quad (\text{D.5})$$

dont la dérivée est donné par

$$\forall w \in \mathbb{R}, \quad f'(w) = (1+w)e^w, \quad (\text{D.6})$$

strictement positive sur $] -1, +\infty[$ et strictement négative sur $]-\infty, -1[$. On en déduit le tableau de variation de f , que l'on complète grâce à

$$\begin{aligned} \lim_{w \rightarrow -\infty} f(w) &= 0, \\ \lim_{w \rightarrow +\infty} f(w) &= +\infty, \\ f(-1) &= -\frac{1}{e}, \\ f(0) &= 0. \end{aligned}$$

x	$-\infty$	-1	0	$+\infty$
signe de $v'(x)$		$-$	0	$+$
variations de v	0	$-\frac{1}{e}$	0	$+\infty$

TABLE D.1. Tableau de variation de f

Voir le tableau D.1 et les figure 1(a) et 1(b).

Compte tenu du tableau D.1, on constate que

— l'image de \mathbb{R} par f est égale à $[-1/e, +\infty[$;

— f est strictement décroissante sur $] -\infty, -1]$ et définit donc une bijection de $] -\infty, -1]$ sur $f(] -\infty, -1]) = [-1/e, 0[$;

— f est strictement croissante sur $[-1, +\infty[$ et définit donc une bijection de $[-1, +\infty[$ sur $f([-1, +\infty[) = [-1/e, +\infty[$.

Voir les figures 1(c) et 1(d). On peut donc conclure à partir de ces éléments. \square

D.2. Utilisation de la fonction W de Lambert : résolution de l'équation $ae^x + bx + c = 0$

LEMME D.2. Soient a, b et c trois réels, a et b étant non nuls. On pose

$$\Delta = \frac{a}{b} e^{-\frac{c}{b}}. \quad (\text{D.7})$$

L'équation

$$ae^x + bx + c = 0, \quad x \in \mathbb{R}, \quad (\text{D.8})$$

admet

- aucune solution si $\Delta < -\frac{1}{e}$;
- une unique solution x si $\Delta = -\frac{1}{e}$, donnée par

$$x = -\frac{c}{b} + 1, \quad (\text{D.9})$$

- deux solutions deux à deux distinctes x_0 et x_{-1} si Δ appartient à $] -\frac{1}{e}, 0[$, données par

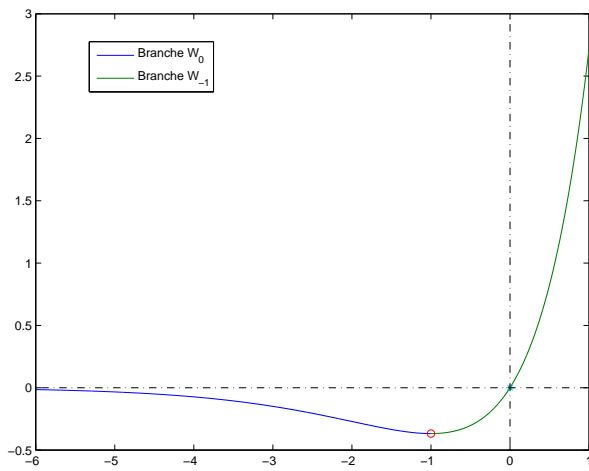
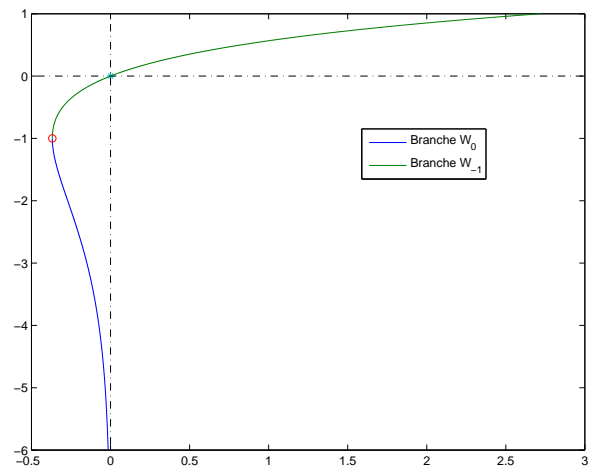
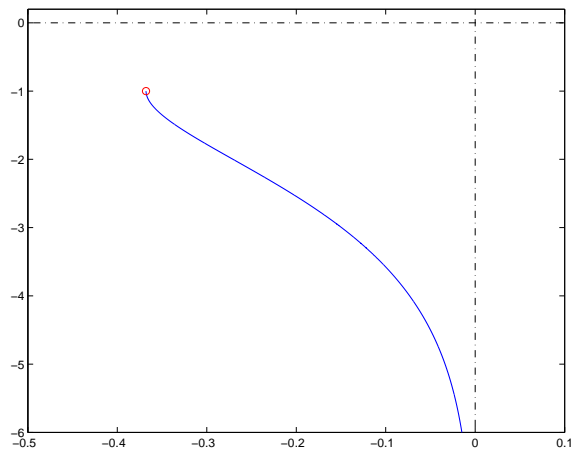
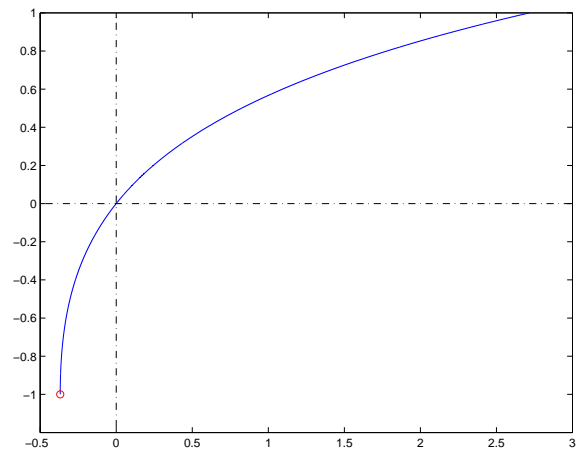
$$x_k = -\frac{c}{b} - W_k(\Delta), \quad \forall k \in \{0, -1\}, \quad (\text{D.10})$$

- une unique solution x si Δ appartient à $[0, +\infty[$, donnée par

$$x = -\frac{c}{b} - W_0(\Delta). \quad (\text{D.11})$$

DÉMONSTRATION. L'équation (D.8) est successivement équivalente à (puisque a et b sont non nuls)

$$\begin{aligned} \frac{a}{b} e^x = -x - \frac{c}{b} &\iff \left(-x - \frac{c}{b}\right) \frac{b}{a} e^{-x} = -1, \\ &\iff \left(-x - \frac{c}{b}\right) e^{-x} e^{-\frac{c}{b}} = \frac{a}{b} e^{-\frac{c}{b}}, \end{aligned}$$

(a) La fonction f et les deux branches W_0 (en vert) et W_{-1} (en bleu).(b) La fonction f^{-1} et les deux branches W_0 (en vert) et W_{-1} (en bleu).(c) la branche W_{-1} .(d) la branche W_0 .FIGURE D.1. La fonction f et les deux branches W_0 et W_{-1} .

cela est équivalent, en posant

$$X = -x - \frac{c}{b},$$

$$\Delta = \frac{a}{b} e^{-\frac{c}{b}},$$

à

$$Xe^X = \Delta. \quad (\text{D.12})$$

Il ne reste plus qu'à appliquer les résultats de la proposition D.1, qui donne X en fonction de $W_k(\Delta)$ et donc x de la forme (selon les valeurs de Δ) :

$$x = -\frac{c}{b} - W_k(\Delta).$$

□

D.3. Utilisation de la fonction W de Lambert : résolution de l'équation $x^x = z$

On cherche à résoudre l'équation suivante : pour $z \in \mathbb{R}$ donné, on cherche $x \in \mathbb{R}$ tel que

$$x^x = z. \quad (\text{D.13})$$

Si x existe, on a nécessairement

$$z > 0. \quad (\text{D.14})$$

L'équation (D.13) est équivalente à

$$e^{x \ln x} = z,$$

soit encore

$$x \ln x = \ln z,$$

et donc

$$\ln x e^{\ln x} = \ln z,$$

et donc encore

$$X e^X = \ln z. \quad (\text{D.15})$$

où

$$X = \ln x. \quad (\text{D.16})$$

D'après la section D.1, cette équation a une solution supérieur à -1 unique, si $\ln(z) \geq -1/e$ et donc si

$$z \geq e^{-\frac{1}{e}}. \quad (\text{D.17})$$

Cette solution est donnée par

$$X = W(\ln z),$$

et, d'après (D.16)

$$x = e^{W(\ln z)},$$

et d'après l'équation (D.1), on a finalement

$$x = \frac{\ln z}{W(\ln z)} \geq \frac{1}{e}. \quad (\text{D.18})$$

Méthode de dichotomie ou de bisection (sous la forme d'un exercice corrigé)

Donnons dans cette annexe, [BM03, Exercice 4.1].

Énoncé

Voir [BM03, TP 4.A]

Soit f une fonction de \mathbb{R} dans \mathbb{R} continue. On suppose $f(a)f(b) \leq 0$; la fonction f admet une racine dans $[a, b]$ dont on cherche une approximation.

- (1) On construit les trois¹ suites $(x_n)_{n \in \mathbb{N}}$, $(a_n)_{n \in \mathbb{N}}$ et $(b_n)_{n \in \mathbb{N}}$ définies de la façon suivante par récurrence : on pose $a_0 = a$, $b_0 = b$ et pour tout $n \in \mathbb{N}$, $x_n = (a_n + b_n)/2$ avec

$$\begin{aligned} \text{si } f(a_n)f(x_n) \leq 0, & \quad a_{n+1} = a_n, \quad b_{n+1} = x_n, \\ \text{si } f(a_n)f(x_n) > 0, & \quad a_{n+1} = x_n, \quad b_{n+1} = b_n. \end{aligned}$$

Vérifier que les trois suites $(x_n)_{n \in \mathbb{N}}$, $(a_n)_{n \in \mathbb{N}}$ et $(b_n)_{n \in \mathbb{N}}$ sont définies et que, pour tout $n \in \mathbb{N}$,

$$f(a_n)f(b_n) \leq 0. \tag{E.1}$$

- (2) Montrer que, pour tout $n \in \mathbb{N}$,

$$|a_{n+1} - b_{n+1}| \leq \frac{1}{2}|a_n - b_n|, \tag{E.2}$$

et

$$a \leq a_n \leq x_n \leq b_n \leq b, \quad a_{n+1} \geq a_n, \quad b_{n+1} \leq b_n. \tag{E.3}$$

- (3) En déduire que, pour tout $n \in \mathbb{N}$,

$$|a_n - b_n| \leq \frac{1}{2^n}(b - a). \tag{E.4}$$

- (4) En déduire que, si f est continue sur $[a, b]$, alors les trois suites $(a_n)_{n \in \mathbb{N}}$, $(b_n)_{n \in \mathbb{N}}$ et $(x_n)_{n \in \mathbb{N}}$ convergent vers une racine de f , notée α . Montrer que, pour tout $n \in \mathbb{N}$,

$$|x_n - \alpha| \leq \frac{1}{2^n}(b - a). \tag{E.5}$$

- (5) Si $\varepsilon > 0$ est donné, déterminer en fonction de ε , le plus petit entier n tel que

$$|x_n - \alpha| \leq \varepsilon. \tag{E.6}$$

- (6) Quel est l'avantage de la méthode de la dichotomie ?

1. Souvent, on ne s'intéresse qu'aux deux suites $(a_n)_{n \in \mathbb{N}}$ et $(b_n)_{n \in \mathbb{N}}$ qui encadrent la racine recherchée. On rajoute ici la suite $(x_n)_{n \in \mathbb{N}}$, par analogie aux autres méthodes qui font intervenir une telle suite.

Corrigé

- (1) On montre aisément par récurrence sur n , que, pour tout $n \in \mathbb{N}$, $f(a_n)f(b_n) \leq 0$, ce qui implique que les suites $(x_n)_{n \in \mathbb{N}}$, $(a_n)_{n \in \mathbb{N}}$ et $(b_n)_{n \in \mathbb{N}}$ sont correctement définies.
- (2) L'inégalité (E.2) est immédiate par récurrence sur n , puisqu'à chaque itération, l'intervalle $[a_n, b_n]$ est découpé en deux (d'où le nom de la méthode²). Quant aux inégalités (E.3), elles résultent de la définition de x_n , a_n et b_n .
- (3) L'égalité (E.4) se démontre à partir de (E.2) par récurrence sur n .
- (4) La suite $(a_n)_{n \in \mathbb{N}}$ (resp. $(b_n)_{n \in \mathbb{N}}$) est donc croissante (resp. décroissante) et majorée par b (resp. minorée par a). Ainsi, elles admettent chacune une limite notée α et α' . D'après (E.4), à la limite $\alpha = \alpha'$. D'autre part, selon (E.1), puisque f est continue, à la limite $f(\alpha)^2 \leq 0$. Ainsi, $f(\alpha) = 0$ et les deux suite $(a_n)_{n \in \mathbb{N}}$ et $(b_n)_{n \in \mathbb{N}}$ convergent vers une racine de f . Ainsi, selon (E.3), la suite $(x_n)_{n \in \mathbb{N}}$ converge vers la même racine, d'après le théorème des gendarmes³.

D'autre part, puisque $(a_n)_{n \in \mathbb{N}}$ est croissante, on a nécessairement, pour tout n , $a_n \leq \alpha$. De même, on a $b \geq \alpha$. Ainsi, pour tout n , $\alpha \in [a_n, b_n]$. Selon (E.3) et (E.4), on en déduit (E.5).

REMARQUE E.1. Attention, f peut avoir plusieurs racines et les suites $(x_n)_{n \in \mathbb{N}}$, $(a_n)_{n \in \mathbb{N}}$ et $(b_n)_{n \in \mathbb{N}}$ peuvent tendre vers l'une ou l'autre de ces racines.

REMARQUE E.2. Les deux suites $(a_n)_{n \in \mathbb{N}}$ et $(b_n)_{n \in \mathbb{N}}$ sont adjacentes : elles vérifient la définition E.3 et le lemme E.4 (pour plus de complément, consulter [RDO88, section 1.2.2. 2°]) :

DÉFINITION E.3. Deux suites de réels $(a_n)_{n \in \mathbb{N}}$ et $(b_n)_{n \in \mathbb{N}}$ sont dites adjacentes si et seulement si la suite $(a_n)_{n \in \mathbb{N}}$ est croissante, la suite $(b_n)_{n \in \mathbb{N}}$ est décroissante et si $(a_n - b_n)_{n \in \mathbb{N}}$ tend vers zéro.

LEMME E.4. *Deux suites adjacentes convergent vers la même limite.*

On peut donc montrer grâce à (E.3) et (E.4) directement que les suites $(a_n)_{n \in \mathbb{N}}$ et $(b_n)_{n \in \mathbb{N}}$ sont adjacentes et en déduire grâce au lemme E.4 leur convergence ainsi que celle x_n , vers la même limite qui est une racine de f .

- (5) Pour que (E.6) soit vrai, il suffit que $1/2^n(b - a) \leq \varepsilon$ ce qui donne (A.26) et (A.27).
- (6) La méthode de dichotomie assure la convergence des suites (a_n) et (b_n) et on est sûr de ne pas «sortir» de l'intervalle initial $[a, b]$ grâce à (E.3).

REMARQUE E.5. On verra plus tard que la convergence (E.5) est «lente». Néanmoins, cette méthode est globalement convergente sur $[a, b]$, puisque $x_n \in [a, b]$. Ainsi, la méthode de la dichotomie permet de s'approcher lentement mais sûrement d'une racine de f . «Une fois assez proche», on pourra alors utiliser d'autres méthodes plus rapides, d'ordre supérieur à un. Leur convergence nécessite en général cette approche préalable.

2. Dichotomie est emprunté en 1750 du grec *dikhotomia*, «division en deux parties égales», élément correspondant à l'adverbe *dikha*, «en deux», et de *-tomia*, «division», «section». cf [Rey98]

3. Les deux gendarmes a_n et b_n encadrent x_n , qui n'a pas d'autre choix que de converger, contraint et forcé, vers la limite commune.

Convergence globale de la méthode du point fixe

En cours de rédaction.

Nous donnons dans cette annexe des résultats de convergence globale de la méthode du point fixe, qui peuvent compléter la proposition A.19.

F.1. Cas particuliers

PROPOSITION F.1. *Fonction g continue ayant un nombre fini de points fixes α_i avec chaque intervalle $[\alpha_i, \alpha_{i+1}]$ g -stables.*

Exemple : résultat relatif à g^2 du cas 2b de l'annexe H.

Dans ce cas, entre chaque point fixe, on a $g(x) > x$ ou $g(x) < x$. De plus, cela implique, avec la g stabilité de $[\alpha_i, \alpha_{i+1}]$ une majoration ou une minoration de la valeur absolue de la dérivée de g par rapport à 1 (et donc point attractif ou répulsif, dont l'aspect seul ne montre pas *a priori* la convergence ou la divergence de la méthode.

PROPOSITION F.2. *Fonction g continue ayant un point fixe unique α appartenant à l'intérieur de l'intervalle I , g -stable, tel que*

$$\begin{aligned}\forall x \in I, x \leq \alpha &\implies g(x) \geq \alpha, \\ x \geq \alpha &\implies g(x) \leq \alpha,\end{aligned}$$

telle que $g^2 = g \circ g$ vérifie les hypothèses de la proposition F.1.

Exemples : Annexes H ou I.

F.2. Cas général

THÉORÈME F.3. *À rédiger*

Étude de la convergence globale de la suite définie par $u_0 \in \mathbb{R}_+^*$ et $u_{n+1} = 1/2(u_n + A/u_n)$, sous la forme d'un problème corrigé

Nous montrons dans cette annexe, la convergence globale sur \mathbb{R}_+^* de la suite définie par $u_{n+1} = 1/2(u_n + A/u_n)$, c'est-à-dire sa convergence pour tout $u_0 \in \mathbb{R}_+^*$.

Énoncé

Pour $A \in \mathbb{R}_+^*$, on considère la suite u_n définie par $u_0 \in \mathbb{R}$ et, pour tout $n \in \mathbb{N}$, $u_{n+1} = 1/2(u_n + A/u_n)$.

- (1) Étudier la convergence de la suite (u_n) pour $u_0 \geq \sqrt{A}$.
- (2) Montrer qu'il y a encore convergence de la suite pour $u_0 \leq \sqrt{A}$.
- (3) En introduisant la suite auxiliaire définie par $w_n = (u_n - \sqrt{A})/(u_n + \sqrt{A})$, montrer que $w_{n+1} = w_n^2$ et Conclure.

Pour $u_0 = 3$ et $A = 3$, comment choisir n pour que $|u_n - \sqrt{A}| \leq \varepsilon = 10^{-9}$? Calculer les valeurs de u_n correspondant et conclure.

Corrigé

On pourra consulter [BM03, Exercice 4.4] qui a inspiré cet exercice.

(1)

Voir, par exemple, la figure G.1.

Commençons par étudier la fonction $g(x) = 1/2(x + A/x)$ sur \mathbb{R}_+^* . On a

$$g'(x) = \frac{1}{2} \left(1 - \frac{A}{x^2} \right) = \frac{1}{2x^2} (x^2 - A),$$

strictement négative sur $]0, \sqrt{A}[$ et strictement positive pour $]\sqrt{A}, +\infty[$. Ainsi g est strictement décroissante sur $]0, \sqrt{A}[$ et strictement croissante pour $]\sqrt{A}, +\infty[$. De plus,

$$\begin{aligned} g(\sqrt{A}) &= \sqrt{A}, \\ \lim_{x \rightarrow 0^+} g(x) &= +\infty, \\ \lim_{x \rightarrow +\infty} g(x) &= +\infty. \end{aligned}$$

Voir le tableau de variation G.1 page suivante. L'asymptote de g en $+\infty$ est la droite d'équation $y = x/2$. On en déduit le tracé de g .

Voir par exemple la figure G.2. On déduit en particulier de l'étude de g que si $J = [\sqrt{A}, +\infty[$, alors J est g -stable (voir définition A.18 page 70).

Remarquons aussi que si on pose $f(x) = g(x) - x$, on a

$$f(x) = \frac{1}{2x} (x^2 + A - 2x^2) = \frac{1}{2x} (A - x^2),$$

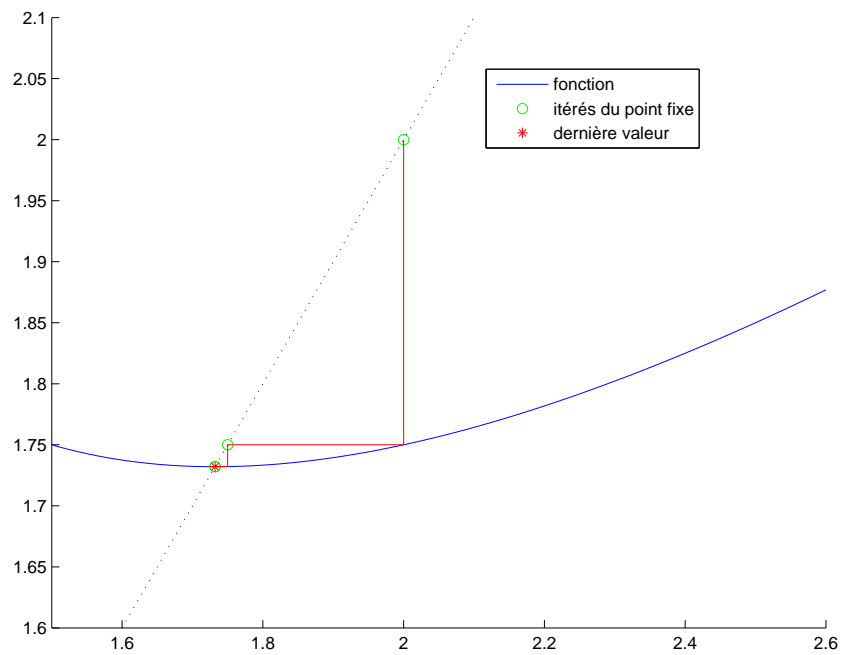


FIGURE G.1. Le tracé graphique des valeurs de u_n pour $u_0 = 2$ et $A = 3$.

x	0	\sqrt{A}	$+\infty$
signe de $g'(x)$		-	+
variations de g	$+\infty$	\sqrt{A}	$+\infty$

TABLE G.1. Tableau de variation de g

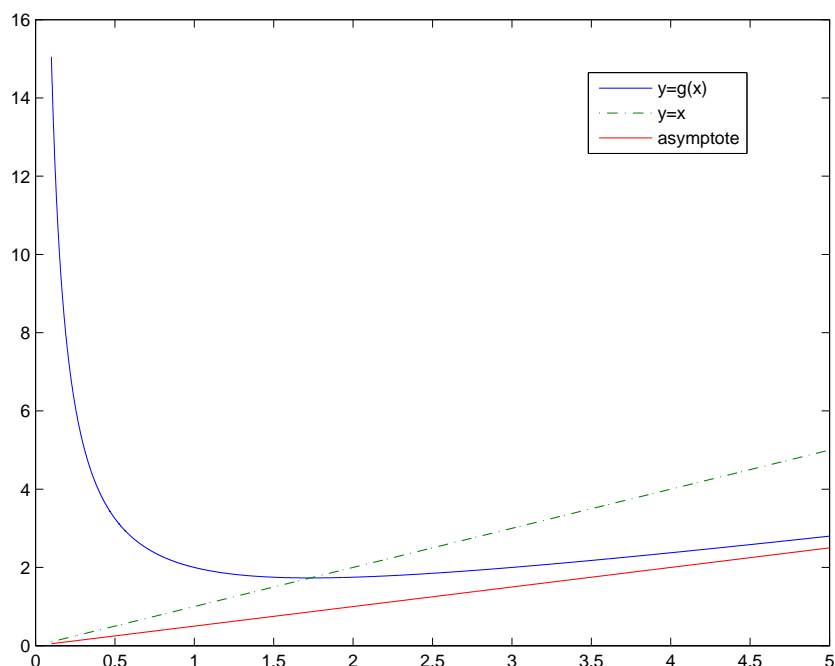
nulle en \sqrt{A} , strictement négative sur $[\sqrt{A}, +\infty[$, strictement positive sur $]\sqrt{A}, +\infty[$. g n'admet donc qu'un seul point fixe, \sqrt{A} . Ainsi, pour tout $x \geq \sqrt{A}$, $g(x) \leq x$. Si $u_0 \in I$, alors d'après ce qui précède, pour tout n , u_n appartient à I et donc, $u_{n+1} \leq u_n$. La suite u_n est décroissante et minorée par \sqrt{A} ; elle est donc convergente vers l'unique point fixe de g , qui est \sqrt{A} .

On peut aussi, de façon alternative, remarquer que pour tout n , $u_n \in I = [\sqrt{A}, u_1]$ pour tout n . On a aussi, pour tout $x \in I$, $0 < g'(x)$ et

$$g'(x) - 1/2 = -\frac{1}{2} \frac{A}{x^2},$$

et, pour tout $x \in I$, on a $1/x^2 \leq A$ et donc

$$g'(x) - 1/2 \geq -\frac{1}{2}.$$

FIGURE G.2. La fonction g pour $A = 3$.

On peut donc appliquer la proposition A.19 page 70, avec $k = 1/2$.

(2)

Voir, par exemple, la figure G.3.

Si $0 < u_0 \leq \sqrt{A}$, on montre par récurrence que pour tout n , $u_n > 0$ et que la suite est correctement définie. De plus, puisque $g(]0, \sqrt{A}) = [\sqrt{A}, +\infty[$, alors u_1 appartient à $[\sqrt{A}, +\infty[$ et il suffit d'appliquer la question 1.

REMARQUE G.1. La convergence de la suite u_n pour toute valeur de u_0 pouvait aussi se déduire du théorème 4 de l'annexe intitulée "Convergence globale de la méthode de Newton" de [DB21].

(3) Par définition,

$$\begin{aligned} w_{n+1} &= \frac{\frac{1}{2} \left(u_n + \frac{A}{u_n} \right) - \sqrt{A}}{\frac{1}{2} \left(u_n + \frac{A}{u_n} \right) + \sqrt{A}}, \\ &= \frac{u_n^2 - 2\sqrt{A}u_n + A}{u_n^2 + 2\sqrt{A}u_n + A}, \\ &= \frac{(u_n - \sqrt{A})^2}{(u_n + \sqrt{A})^2}. \end{aligned}$$

et donc

$$\forall n \in \mathbb{N}, \quad w_{n+1} = w_n^2. \tag{G.1}$$

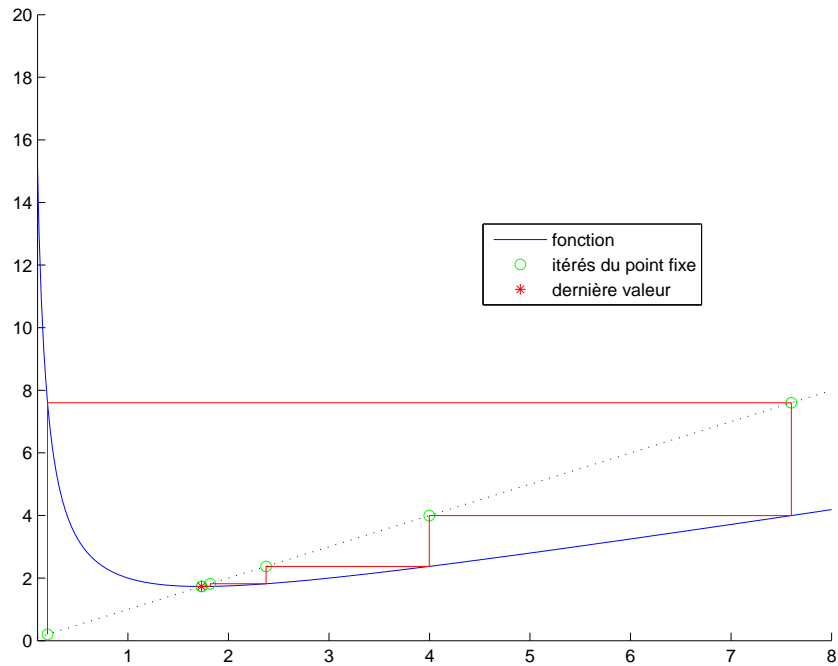


FIGURE G.3. Le tracé graphique des valeurs de u_n pour $u_0 = 0.200$ et $A = 3$.

On retrouve donc l'ordre deux de la convergence de la méthode de Newton. En effet, on peut réécrire (G.1) :

$$\frac{|u_n - \sqrt{A}|}{|u_{n+1} - \sqrt{A}|^2} = \frac{|u_n + \sqrt{A}|}{|u_{n+1} + \sqrt{A}|^2},$$

et donc

$$\lim_{n \rightarrow +\infty} \frac{|u_n - \sqrt{A}|}{|u_{n+1} - \sqrt{A}|^2} = \frac{1}{2\sqrt{A}}.$$

On déduit aussi de (G.1) par récurrence sur n que

$$\forall n \in \mathbb{N}, \quad w_n = w_0^{(2^n)} = \left(\frac{u_0 - \sqrt{A}}{u_0 + \sqrt{A}} \right)^{(2^n)}. \tag{G.2}$$

Puisque le nombre $\frac{u_0 - \sqrt{A}}{u_0 + \sqrt{A}}$ est dans $[0, 1[$, on en déduit que $w_n \rightarrow 0$ quand n tend vers l'infini. En revenant à la définition de w_n :

$$w_n = \frac{u_n - \sqrt{A}}{u_n + \sqrt{A}},$$

on a

$$u_n = (u_n + \sqrt{A}) w_n + \sqrt{A},$$

et puisque u_n est bornée, cela nous permet de retrouver que u_n tend vers \sqrt{A} quand n tend vers l'infini.

(4) On peut déduire de (G.2) que, puisque $(u_n)_{n \in \mathbb{N}}$ est décroissante,

$$|u_n - \sqrt{A}| \leq (u_n + \sqrt{A}) w_n \leq (L_0 + \sqrt{A}) \left(\frac{L_0 - \sqrt{A}}{L_0 + \sqrt{A}} \right)^{(2^n)}.$$

Ainsi, pour que $|u_n - \sqrt{A}| \leq \varepsilon$, il suffit que

$$(L_0 + \sqrt{A}) \left(\frac{L_0 - \sqrt{A}}{L_0 + \sqrt{A}} \right)^{(2^n)} \leq \varepsilon, \quad (\text{G.3})$$

donc que

$$(L_0 + 2) \left(\frac{L_0}{L_0 + 1} \right)^{(2^n)} \leq \varepsilon,$$

car $u_0 = 3$ et $A = 3$. Soit encore

$$n \geq \frac{\ln \left(\frac{\ln(\varepsilon/5)}{\ln(3/4)} \right)}{\ln 2}.$$

Numériquement, pour $\varepsilon = 10^{-9}$, on trouve $n \geq 7$. Indiquons les différentes valeurs de u_n et de $u_n - \sqrt{A}$, calculées sous matlab :

n	u_n	$ u_n - \sqrt{A} $
0	3	1.26794
1	2	0.26794
2	1.75	0.01795
3	1.73214285714286	9.20496×10^{-5}
4	1.73205081001473	2.44585×10^{-9}
5	1.73205080756888	0
6	1.73205080756888	0
7	1.73205080756888	0

Les résultats sont bien conformes à ce que on avait prévu (mais, comme souvent, la majoration était pessimiste, en raison du caractère grossier de la majoration utilisée).

Étude du zéro de la fonction $e^x + x - K$, sous la forme d'un problème corrigé

Cet exercice a été donné partiellement à l'examen d'Automne 2020

Énoncé

Soit $K \in \mathbb{R}$. Dans cet exercice, on étudie l'équation

$$e^x = -x + K. \quad (\text{H.1})$$

- (1) Montrer que l'équation (H.1) admet une unique solution sur \mathbb{R} , notée r .
 (2) On définit la fonction g par

$$\forall x \in \mathbb{R}, \quad g(x) = K - e^x, \quad (\text{H.2})$$

On met désormais l'équation (H.1) sous la forme

$$g(x) = x, \quad (\text{H.3})$$

et on considère la méthode du point fixe associée sur un intervalle $[a, b]$ (définie par $x_0 \in [a, b]$ et $x_{n+1} = g(x_n)$).

- (a) On pose, dans cette question :

$$a = -2, \quad (\text{H.4a})$$

$$b = -1/2, \quad (\text{H.4b})$$

$$K = -1. \quad (\text{H.4c})$$

- (i) Montrer qu'avec ces valeurs la méthode du point fixe converge pour tout $x_0 \in [a, b]$.
 (ii) Déterminer la valeur de n telle que $|x_n - r| \leq \varepsilon$ avec

$$\varepsilon = 10^{-2}. \quad (\text{H.5})$$

- (iii) Pour cette valeurs de n , calculer les valeurs de $(x_p)_{0 \leq p \leq n}$ avec $x_0 = -5/4$.

- (b) On pose, dans cette question :

$$a = 1/4, \quad (\text{H.6a})$$

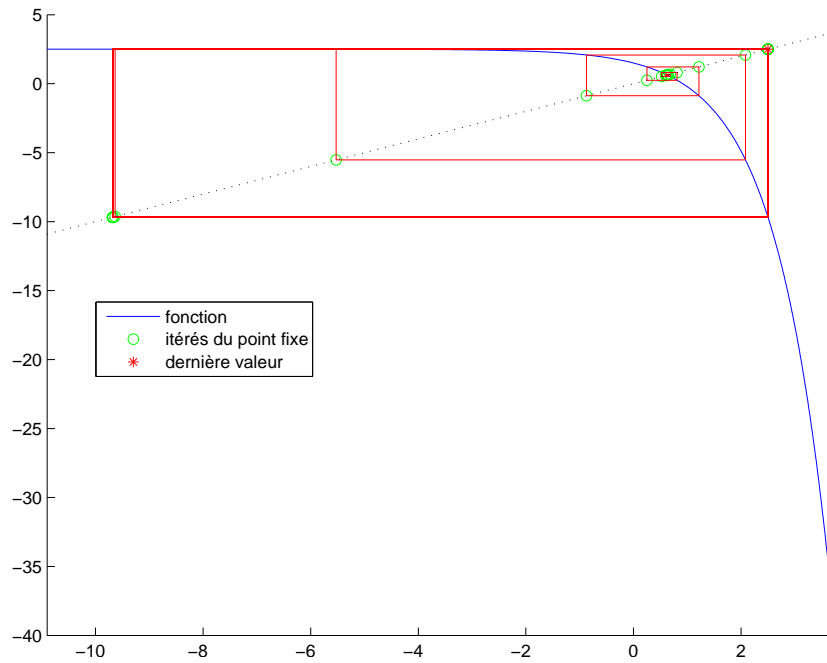
$$b = 1, \quad (\text{H.6b})$$

$$K = 5/2. \quad (\text{H.6c})$$

- (i) Est-ce que la méthode du point fixe converge?
 (ii)

On a affiché sur la figure H.1 page suivante, les 30 premières valeurs calculées avec les paramètres donnés par (H.6) et x_0 donné par

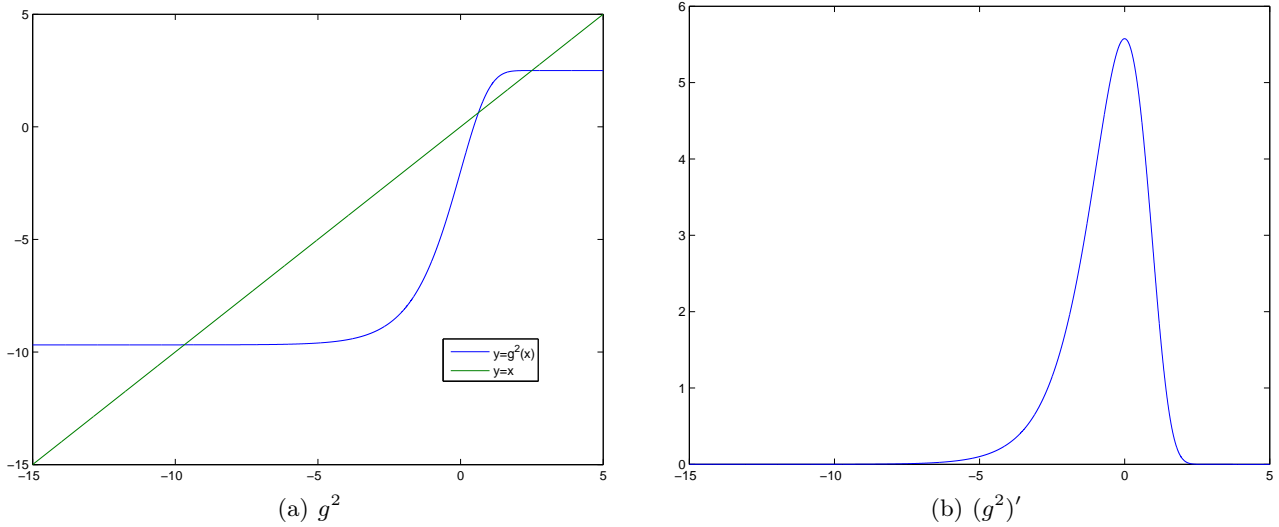
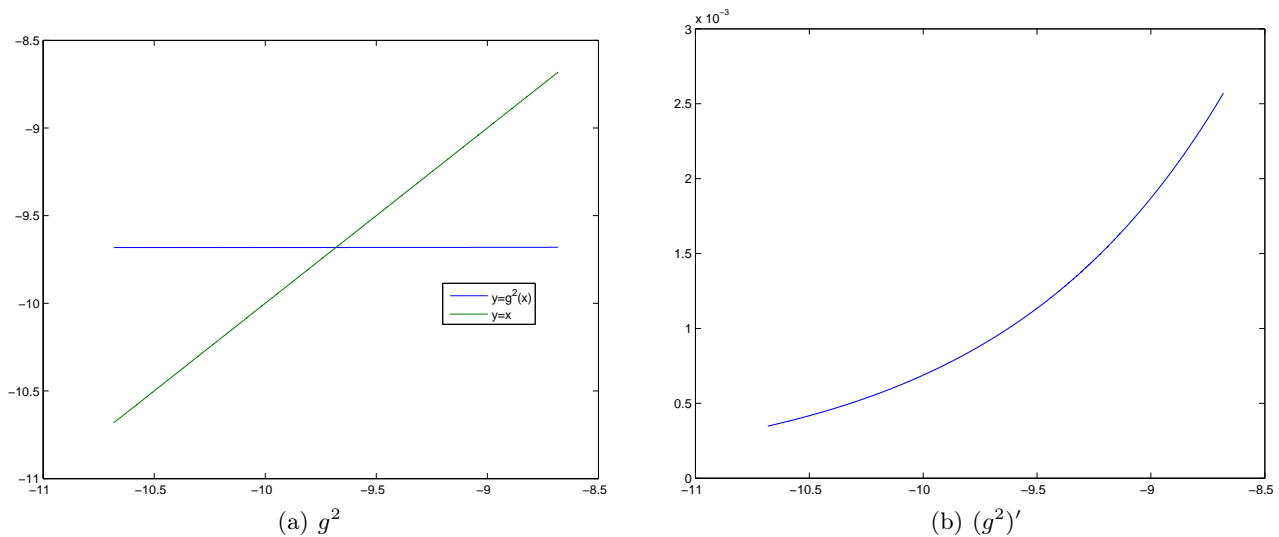
$$x_0 = 5/8. \quad (\text{H.7})$$

FIGURE H.1. Les 30 premières valeurs de x_n pour $x_0 = 5/8$ et $K = 5/2$.

n	x_{2n}	x_{2n+1}
0	0.625000000000000	0.63175404256778
1	0.61909312177470	0.64275701533088
2	0.59828327950539	0.68100658482373
3	0.52413439208549	0.81100379305089
4	0.24983444603184	1.21618714121982
5	-0.87429745515325	2.08284501315852
6	-5.52727415891270	2.49602318550744
7	-9.63414264798875	2.49993454467100
8	-9.68169657765000	2.49993758447888
9	-9.68173360772356	2.49993758679009
10	-9.68173363587809	2.49993758679185
11	-9.68173363589949	2.49993758679185
12	-9.68173363589951	2.49993758679185
13	-9.68173363589951	2.49993758679185
14	-9.68173363589951	2.49993758679185

TABLE H.1. Les 30 premières valeurs de x_n pour $x_0 = 5/8$ et $K = 5/2$.

On a indiqué dans le tableau H.1, les valeurs correspondantes, en séparant les termes d'indices impairs et pairs.

FIGURE H.2. Les graphiques des fonction g^2 et $(g^2)'$ sur l'intervalle $[-15, 5]$.FIGURE H.3. Les graphiques des fonction g^2 et $(g^2)'$ sur l'intervalle $[-10.68173, -8.68173]$.

(A)

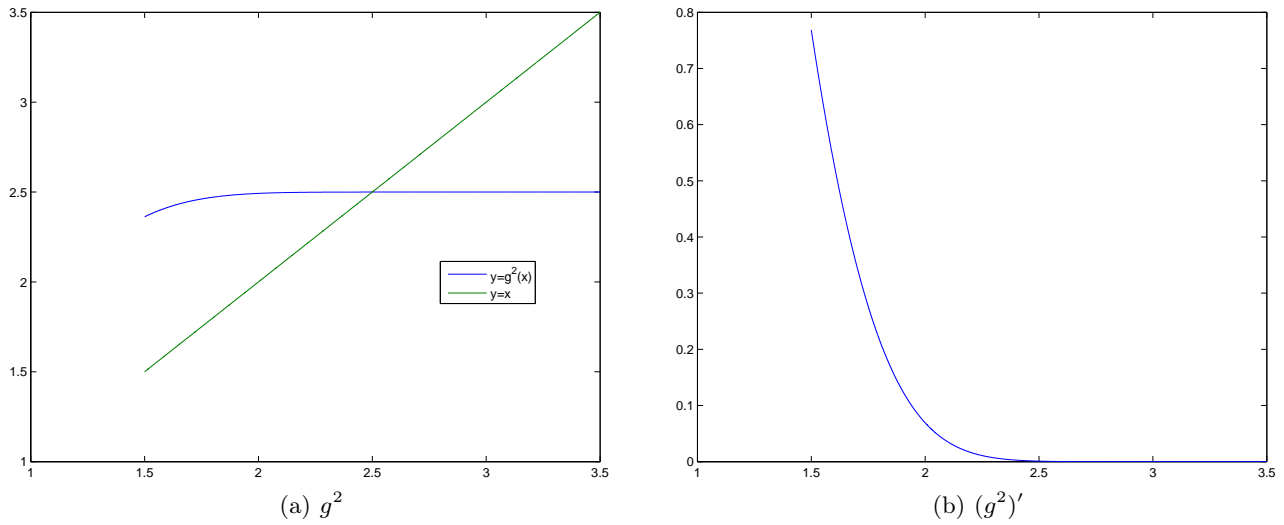
Que constatez-vous sur cette figure et sur ce tableau ?

(B)

On définit la fonction g^2 par

$$\forall x \in \mathbb{R}, \quad g^2(x) = g(g(x)). \quad (\text{H.8})$$

On a tracé sur la figure H.2, les graphes des fonctions g^2 et $(g^2)'$ sur l'intervalle $[-15, 5]$, sur la figure H.3, les graphes des fonctions g^2 et $(g^2)'$ sur l'intervalle $[-10.68173, -8.68173]$ et sur la figure H.4 page suivante les graphes des fonctions g^2 et $(g^2)'$ sur l'intervalle $[1.49994, 3.49994]$. Au vu de ces figures, essayez d'expliquer les observations faites dans la question 2(b)iiA.

FIGURE H.4. Les graphiques des fonction g^2 et $(g^2)'$ sur l'intervalle $[1.49994, 3.49994]$.

(iii) Quelle méthode pourriez-vous utiliser pour résoudre (H.1) dans ce cas ?

Corrigé

(1) Soit $K \in \mathbb{R}$. Montrons que l'équation

$$e^x = -x + K. \quad (\text{H.9})$$

admet une unique solution sur \mathbb{R} , notée r . Pour cela, on définit la fonction f sur \mathbb{R} par

$$\forall x \in \mathbb{R}, \quad f(x) = e^x + x - K, \quad (\text{H.10})$$

dont les zéros sont exactement les solutions de (H.9). La fonction f est dérivable sur \mathbb{R} et on a

$$\forall x \in \mathbb{R}, \quad f'(x) = e^x + 1, \quad (\text{H.11})$$

qui est strictement positive. La fonction f est donc strictement croissante sur \mathbb{R} . De plus, on a

$$\lim_{x \rightarrow -\infty} f(x) = -\infty, \quad (\text{H.12a})$$

$$\lim_{x \rightarrow +\infty} f(x) = +\infty, \quad (\text{H.12b})$$

dont on déduit, d'après la continuité de f et le théorème des valeurs intermédiaires que

$$f \text{ admet un unique zéro sur } \mathbb{R}, \text{ noté } r \text{ (ou } r(K) \text{ en cas d'ambiguïté)}. \quad (\text{H.13})$$

Par définition, on a donc

$$e^{r(K)} + r(K) - K = 0. \quad (\text{H.14})$$

REMARQUE H.1. Remarquons que la stricte croissance de f sur \mathbb{R} implique, puisque r est le zéro de f :

$$\forall x \in \mathbb{R}, \quad x > r \implies f(x) > f(r) = 0,$$

$$x < r \implies f(x) < 0,$$

$$x = r \implies f(x) = 0,$$

et donc, puisque

$$\forall x \in \mathbb{R}, \quad f(x) = x - g(x), \quad (\text{H.15})$$

on obtient naturellement que

$$g \text{ n'a qu'un seul point fixe sur } \mathbb{R}, \text{ qui est } r \text{ (noté } r(K) \text{ en cas d'ambiguïté)}. \quad (\text{H.16})$$

et de plus

$$\forall x \in \mathbb{R}, \quad x > r \implies g(x) < x, \quad (\text{H.17a})$$

$$x < r \implies g(x) > x, \quad (\text{H.17b})$$

$$x = r \implies g(x) = x. \quad (\text{H.17c})$$

REMARQUE H.2. On a aussi

$$\forall K < 1, \quad |g'(r(K))| < 1, \quad (\text{H.18a})$$

$$\forall K > 1, \quad |g'(r(K))| > 1, \quad (\text{H.18b})$$

$$g'(r(1)) = -1. \quad (\text{H.18c})$$

En effet, on a

$$\forall x \in \mathbb{R}, \quad g'(x) = -e^x, \quad (\text{H.19})$$

et donc

$$\forall K \in \mathbb{R}, \quad g'(r(K)) = -e^{r(K)}.$$

Remarquons que d'après la stricte croissance de f , on a

$$r(K) < 0 \iff f(r(K)) < f(0) \iff 0 < 1 - K$$

et donc

$$r(K) < 0 \iff K < 1, \quad (\text{H.20})$$

ce qui montre (H.18a) et (H.18b). Enfin, il est clair que

$$r(1) = 0, \quad (\text{H.21})$$

ce qui montre (H.18c).

(2) On définit la fonction g par

$$\forall x \in \mathbb{R}, \quad g(x) = K - e^x, \quad (\text{H.22})$$

et on met désormais l'équation (H.9) sous la forme

$$g(x) = x, \quad (\text{H.23})$$

et on considère la méthode du point fixe associée sur un intervalle $[a, b]$, définie par

$$x_0 \in [a, b] \text{ et } x_{n+1} = g(x_n). \quad (\text{H.24})$$

REMARQUE H.3. On peut en fait, directement étudier la convergence ou la divergence de la méthode du point fixe (selon les valeurs de K) sans utiliser le théorème A.17, comme le suggérait la suite de l'énoncé (et dont nous donnerons la solution plus bas, c'est-à-dire à partir du point 2a page 118). Pour cela, on utilise des techniques proches de celles utilisées dans les annexes G et I.

La preuve se fait en plusieurs points.

(a) Commençons par définir la fonction g^2 par

$$g^2(x) = g(g(x)) = K - e^{K - e^x}. \quad (\text{H.25})$$

et considérons f_2 définie par (de façon analogue à (H.15))

$$f_2(x) = x - g^2(x) = x - K + e^{K - e^x}. \quad (\text{H.26})$$

On a donc

$$(g^2)'(x) = e^{x + K - e^x}, \quad (\text{H.27})$$

et donc

$$\forall x \in \mathbb{R}, \quad (g^2)'(x) > 0. \quad (\text{H.28})$$

On déduit de (H.28) que

$$g^2 \text{ est strictement croissante sur } \mathbb{R}. \tag{H.29}$$

et que

$$f'_2(x) = 1 - (e^x e^{K-e^x}),$$

et donc

$$f'_2(x) = 1 - e^{x+K-e^x}, \tag{H.30}$$

et donc

$$f'_2(x) > 0 \text{ ssi } x + K - e^x < 0,$$

ce qui est équivalent à

$$f'_2(x) > 0 \text{ ssi } h(x) > 0, \text{ où } \forall x \in \mathbb{R}, \quad h(x) = e^x - x - K \tag{H.31}$$

Étudions la fonction h . On a

$$\forall x \in \mathbb{R}, \quad h'(x) = e^x - 1,$$

qui est strictement positive sur \mathbb{R}_+^* et strictement négative sur \mathbb{R}_-^* .

x	$-\infty$	0	$+\infty$
signe de $h'(x)$	-	0	+
variations de h	$+\infty$	$1 - K$	$+\infty$

TABLE H.2. Tableau de variation de h .

Voir le tableau de variation de h dans le tableau H.2. La valeur minimale de h est donc donnée par

$$h_{\min} = h(0) = 1 - K.$$

Nous avons alors deux cas :

(i)

x	$-\infty$	0	$+\infty$
signe de $h'(x)$	-	0	+
variations de h	$+\infty$	$1 - K \geq 0$	$+\infty$

TABLE H.3. Tableau de variation de h dans le cas où $K \leq 1$.

Premier cas : on a

$$K \leq 1, \tag{H.32}$$

ce qui est équivalent à $h_{\min} \geq 0$ et donc, d'après le tableau H.3, h est strictement positive sur \mathbb{R}^* , et ainsi, d'après (H.31),

$$f_2 \text{ est strictement croissante sur } \mathbb{R}. \quad (\text{H.33})$$

Enfin, puisque

$$\lim_{x \rightarrow -\infty} f_2(x) = -\infty, \quad (\text{H.34a})$$

$$\lim_{x \rightarrow +\infty} f_2(x) = +\infty, \quad (\text{H.34b})$$

on en déduit que f_2 admet un zéro unique sur \mathbb{R} . Or, d'après (H.16), on sait que

$$\text{le point fixe } r(K) \text{ de } g \text{ est aussi point fixe de } g^2, \quad (\text{H.35})$$

(puisque $g^2(r) = g(r) = r$) et donc

$$\text{le point fixe } r(K) \text{ de } g \text{ est aussi zéro de } f_2, \quad (\text{H.36})$$

et donc ici

$$\text{le point fixe } r(K) \text{ de } g \text{ est aussi l'unique zéro de } f_2. \quad (\text{H.37})$$

On aboutit à des conclusions analogues à celles de la remarque H.1 : la stricte croissance de f_2 sur \mathbb{R} implique, puisque $r(K)$ est le zéro de f_2 :

$$\forall x \in \mathbb{R}, \quad x > r(K) \implies f_2(x) > f_2(r) = 0, \quad (\text{H.38a})$$

$$x < r(K) \implies f_2(x) < 0, \quad (\text{H.38b})$$

$$x = r(K) \implies f_2(x) = 0, \quad (\text{H.38c})$$

et donc, d'après (H.26), on obtient naturellement que

$$g^2 \text{ n'a qu'un seul point fixe sur } \mathbb{R}, \text{ qui est } r \text{ (noté } r(K) \text{ en cas d'ambiguïté)}. \quad (\text{H.39})$$

et de plus

$$\forall x \in \mathbb{R}, \quad x > r(K) \implies g^2(x) < x, \quad (\text{H.40a})$$

$$x < r(K) \implies g^2(x) > x, \quad (\text{H.40b})$$

$$g^2(r(K)) = r(K). \quad (\text{H.40c})$$

REMARQUE H.4. On a aussi

$$\forall K < 1, \quad \left| (g^2)'(r(K)) \right| < 1, \quad (\text{H.41})$$

et

$$(g^2)'(r(1)) = 1. \quad (\text{H.42})$$

On a effet d'après (H.27),

$$(g^2)'(r(K)) = e^{r(K)+K} - e^{r(K)}$$

et la définition (H.14), de $r(K)$ implique

$$\forall K \in \mathbb{R}, \quad (g^2)'(r(K)) = e^{2r(K)}. \quad (\text{H.43})$$

Enfin, (H.20) et (H.43) impliquent (H.41) et (H.42) vient de (H.21).

(ii)

Second cas : on a

$$K > 1. \quad (\text{H.44})$$

Dans ce sous-cas, $h_{\min} = 1 - K > 0$ et la fonction h s'annule en deux réels δ et η , vérifiant $\delta < 0 < \eta$. Voir le tableau de variation H.4 dont on déduit que h est strictement positif sur $] -\infty, \delta[\cup]\eta, +\infty[$ et strictement négatif sur $]\delta, \eta[$, dont on déduit d'après (H.31) que f_2 est strictement croissante sur $] -\infty, \delta] \cup]\eta, +\infty[$ et strictement décroissante sur $]\delta, \eta[$. On rappelle aussi que l'on a (H.34).

On en déduit le tableau de variation H.5.

Par ailleurs, remarquons que

$$f_2(0) = e^{K-1} - K, \quad (\text{H.45a})$$

$$f_2(\ln K) = 1 + \ln K - K. \quad (\text{H.45b})$$

$$(\text{H.45c})$$

x	$-\infty$	δ	0	η	$+\infty$
signe de $h'(x)$		-	0	+	
variations de h	$+\infty$	\searrow	0	\swarrow	$+\infty$
			$1 - K < 0$		

TABLE H.4. Tableau de variation de h dans le cas où $K > 1$.

x	$-\infty$	δ	0	$\ln K$	η	$+\infty$
signe de $f_2'(x)$		+	0	-	0	+
variations de f_2	$-\infty$	\nearrow	$f_2(\delta)$	\searrow	$f_2(\eta)$	$+\infty$
			$f_2(0) > 0$	$f_2(\ln K) < 0$		

TABLE H.5. Tableau de variation de f_2 dans le cas où $K > 1$.

La fonction logarithme étant concave sur \mathbb{R}_+^* , elle est en dessous de sa tangente en tout point et en particulier au point 1, on a donc

$$\forall h > -1, \quad \ln(1 + h) < h,$$

résultat qui peut aussi être obtenu par une étude de fonction ou par une formule de Taylor-Lagrange à l'ordre 2. Si on pose $1 + h = u$ où $u > 0$, on en déduit

$$\forall u > 0, \quad \ln(u) < u - 1,$$

et en particulier, en $K > 0$, on a donc

$$\forall K > 0, \quad 1 + \ln(K) - K < 0, \tag{H.46}$$

ce qui nous prouve, grâce à (H.45b) :

$$\forall K > 0, \quad f_2(\ln K) < 0. \tag{H.47}$$

De la même façon, la fonction exponentielle étant convexe sur \mathbb{R} , elle est au-dessus de sa tangente en tout point et en particulier au point 0, on a donc

$$\forall h \neq 0, \quad e^h > h + 1,$$

résultat qui peut aussi être obtenu par une étude de fonction ou par une formule de Taylor-Lagrange à l'ordre 2. Si on pose $1 + h = u$ où $u \in \mathbb{R} \setminus \{1\}$, on en déduit

$$\forall u \neq 1, \quad e^{u-1} > h,$$

et en particulier, en $K \neq 0$, on a donc

$$\forall K \neq 0, \quad e^{K-1} > K \tag{H.48}$$

ce qui nous prouve, grâce à (H.45a) :

$$\forall K \neq 0, \quad f_2(0) > 0. \quad (\text{H.49})$$

De cela et du tableau de variation H.5, on déduit que (car $\delta < 0 < \eta$),

$$f_2 \text{ ne s'anulle qu'une fois sur } \mathbb{R}_*^-. \quad (\text{H.50})$$

Enfin, remarquons que $\eta > \ln K$, car c'est équivalent (car $\ln K > 0$) à $h(\eta) > h(\ln K)$ soit à $0 > -\ln K$, ce qui est vrai car $K > 1$. D'après (H.47), on peut donc compléter le tableau de variation H.5, dont on déduit que

$$f_2 \text{ ne s'anulle qu'une fois sur }]0, \ln K[\text{ et une fois sur }]\ln(K), +\infty[. \quad (\text{H.51})$$

Autrement dit, d'après (H.50) et (H.51), que f_2 possède trois zéros, deux à deux distincts, le premier dans l'intervalle $] -\infty, 0[$, le deuxième dans l'intervalle $]0, \ln K[$, et le troisième dans l'intervalle $]\ln(K), +\infty[$. Notons que l'unique zéro $r(k)$ de f est aussi dans l'intervalle $]0, \ln K[$. En effet, cela est équivalent à

$$f(0)f(\ln K) < 0,$$

ce qui est équivalent, compte tenu de la définition de f à

$$(1 - K)(K + \ln K - K) < 0,$$

soit encore à

$$(1 - K) \ln K < 0,$$

ce qui est vrai d'après (H.44). D'après (H.36), et par unicité du zéro de f dans $]0, \ln K[$, on en déduit que le deuxième zéro de f_2 (dans l'intervalle $]0, \ln K[$) ne peut être que $r(K)$ et d'après ce qui précède :

f_2 possède trois zéros, deux à deux distincts, le premier, noté $\alpha(K)$, dans l'intervalle $] -\infty, 0[$,

le second, égal à $r(K)$ dans l'intervalle $]0, \ln K[$, et le troisième, noté $\beta(K)$, dans l'intervalle $]\ln(K), +\infty[$. (H.52)

On en déduit le tableau H.6. En revenant à la la fonction g^2 , on en déduit que

g^2 possède trois points fixes, deux à deux distincts, le premier, noté $\alpha(K)$, dans l'intervalle $] -\infty, 0[$,

le second, égal à $r(K)$ dans l'intervalle $]0, \ln K[$, et le troisième, noté $\beta(K)$, dans l'intervalle $]\ln(K), +\infty[$. (H.53)

x	$-\infty$	$\alpha(K)$	δ	$r(K)$	η	$\beta(K)$	$+\infty$
signe de $f_2'(x)$		+	0	-	0	+	
variations de f_2	$-\infty$	↘ 0 ↗	$f_2(\delta)$	↘ 0 ↗	$f_2(\eta)$	↘ 0 ↗	$+\infty$

TABLE H.6. Tableau de variation de f_2 dans le cas où $K > 1$.

On aboutit à des conclusions analogues à celles des équations (H.38) et (H.40) : le tableau de variation H.6 implique

$$\forall x \in \mathbb{R}, \quad x < \alpha(K) \implies f_2(x) < 0, \quad (\text{H.54a})$$

$$\alpha(K) < x < r(K) \implies f_2(x) > 0, \quad (\text{H.54b})$$

$$r(K) < x < \beta(K) \implies f_2(x) < 0, \quad (\text{H.54c})$$

$$x > r(K) \implies f_2(x) > 0, \quad (\text{H.54d})$$

$$x \in \{\alpha(K), r(K), \beta(K)\} \implies f_2(x) = 0, \quad (\text{H.54e})$$

et

$$\forall x \in \mathbb{R}, \quad x < \alpha(K) \implies g^2(x) > x, \quad (\text{H.55a})$$

$$\alpha(K) < x < r(K) \implies g^2(x) < x, \quad (\text{H.55b})$$

$$r(K) < x < \beta(K) \implies g^2(x) > x, \quad (\text{H.55c})$$

$$x > r(K) \implies g^2(x) < x, \quad (\text{H.55d})$$

$$x \in \{\alpha(K), r(K), \beta(K)\} \implies g^2(x) = x, \quad (\text{H.55e})$$

REMARQUE H.5. Notons que l'on a dans ce cas :

$$\left| (g^2)'(\alpha(K)) \right| < 1, \quad (\text{H.56a})$$

$$\left| (g^2)'(\beta(K)) \right| < 1, \quad (\text{H.56b})$$

$$\left| (g^2)'(r(K)) \right| > 1. \quad (\text{H.56c})$$

— La preuve de (H.56c) se fait comme celle de (H.18b) en utilisant (H.20) et (H.43).

— Montrons (H.56b). Pour cela, nous allons utiliser la convexité de g^2 . D'après (H.27), on a

$$(g^2)''(x) = (1 - e^x)e^{x+K-e^x},$$

et donc

$$\forall x > 0, \quad (g^2)''(x) < 0, \quad (\text{H.57a})$$

$$\forall x < 0, \quad (g^2)''(x) > 0. \quad (\text{H.57b})$$

Or, on a $0 < r(K) < \beta(K)$ donc $g(g^2)''$ est strictement négative sur l'intervalle $[r(K), \beta(K)]$ et donc g^2 est strictement concave, c'est-à-dire, $-g^2$ strictement convexe sur l'intervalle $[r(K), \beta(K)]$. On renvoie à [RDO88, section 4.5.1] Cette convexité implique que la dérivée de $-g^2$ en $\beta(K)$ est strictement supérieure à la pente de $-g^2$ entre les points $r(K)$ et $\beta(K)$:

$$(-g^2)'(\beta(K)) > \frac{-g^2(\beta(K)) + g^2(r(K))}{\beta(K) - r(K)},$$

ce qui implique

$$(-g^2)'(\beta(K)) > \frac{-\beta(K) + r(K)}{\beta(K) - r(K)} = -1,$$

et donc

$$(g^2)'(\beta(K)) < 1.$$

Ainsi, compte tenu de (H.28), on a

$$0 < (g^2)'(\beta(K)) < 1,$$

ce qui montre (H.56b).

Illustrons cela en traçant les divers éléments calculés sur la figure H.5.

— Montrons (H.56a). Cette fois, nous n'utiliserons pas la convexité de g^2 sur la totalité de l'intervalle $[\alpha(K), r(K)]$. Mais, on a $\alpha(K) < 0$ donc d'après (H.57b), g^2 est strictement convexe sur l'intervalle $[\alpha(K), \alpha(K)/2]$ et comme précédemment, on en déduit

$$0 < (g^2)'(\alpha(K)) < \frac{g^2(\alpha(K)/2) - g^2(\alpha(K))}{\alpha(K)/2 - \alpha(K)}. \quad (\text{H.58})$$

Or, on a $\alpha(K)/2 \in]\alpha(K), r(K)[$ et d'après (H.55b), on a $g^2(\alpha(K)/2) < \alpha(K)/2$ et d'après (H.58), on a

$$0 < (g^2)'(\alpha(K)) < \frac{\alpha(K)/2 - g^2(\alpha(K))}{\alpha(K)/2 - \alpha(K)} = \frac{\alpha(K)/2 - \alpha(K)}{\alpha(K)/2 - \alpha(K)} = 1,$$

ce qui nous permet de conclure.

Illustrons cela en traçant les divers éléments calculés sur la figure H.6.

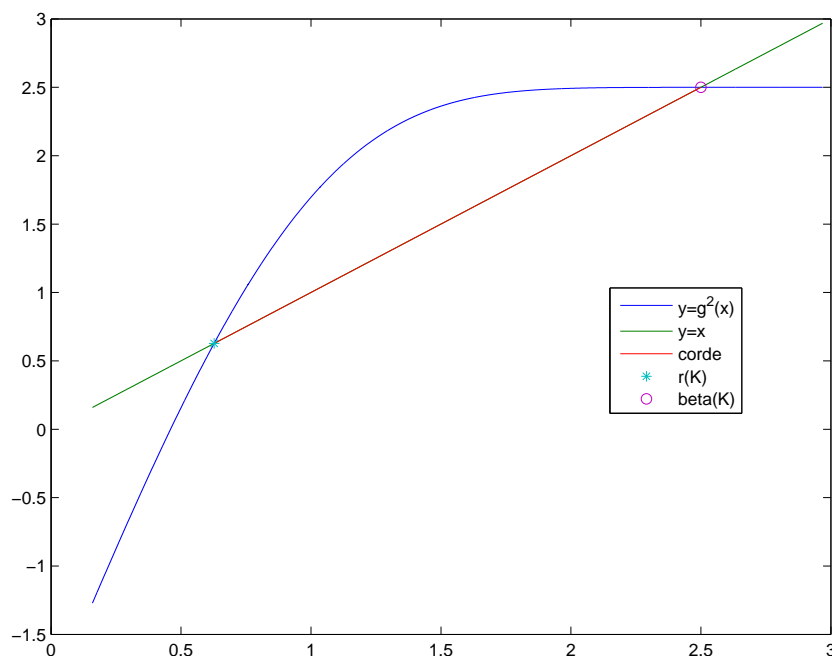


FIGURE H.5. La fonction g^2 sur l'intervalle $[r(K), \beta(K)]$ pour $K = 5/2$.

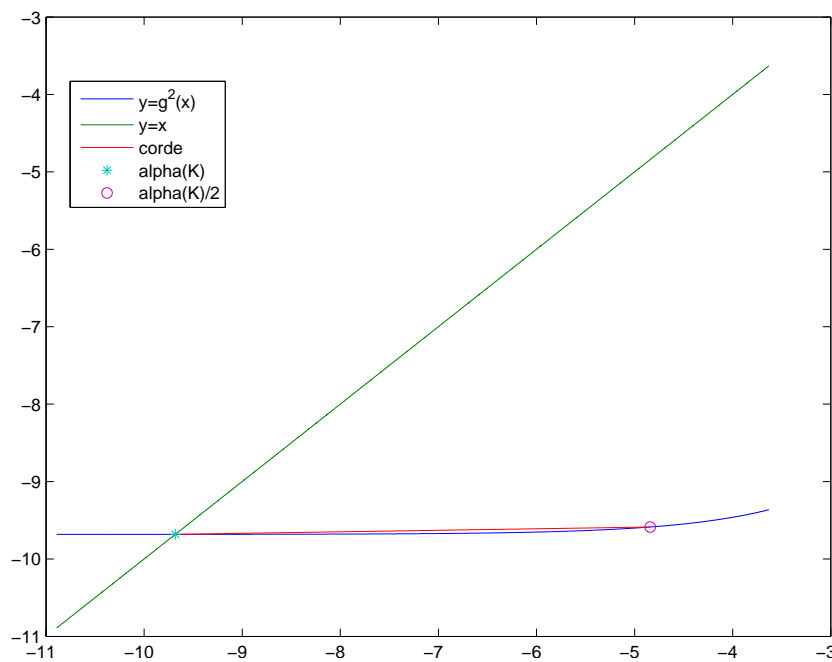


FIGURE H.6. La fonction g^2 sur l'intervalle $[\alpha, \alpha/2]$ pour $K = 5/2$.

Sur la figure (H.7), on constate que $(g^2)'(\alpha(K))$ et $(g^2)'(\beta(K))$ semblent appartenir à $]0, 1[$, pour tout K , ce qui est confirmé numériquement par

$$\begin{aligned} \min_{K \in]1, 15]} (g^2)'(\alpha(K)) &= 0, \\ \max_{K \in]1, 15]} (g^2)'(\alpha(K)) &= 0.997202351770, \\ \min_{K \in]1, 15]} (g^2)'(\beta(K)) &= 0, \\ \max_{K \in]1, 15]} (g^2)'(\beta(K)) &= 0.997202351770. \end{aligned}$$

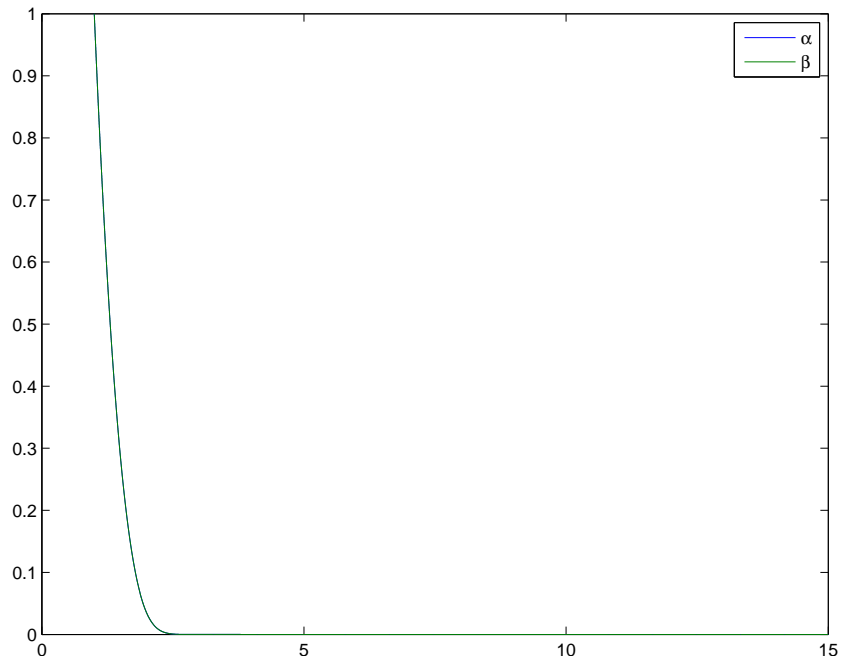


FIGURE H.7. Les courbes $(g^2)'(\alpha(K))$ et $(g^2)'(\beta(K))$ en fonction de K pour $K \in]1, 15]$.

De plus, il semblerait sur cette figure que

$$\forall K, \quad (g^2)'(\alpha(K)) = (g^2)'(\beta(K)),$$

ce qui est confirmé numériquement par

$$\max_{K \in]1, 15]} \left| (g^2)'(\alpha(K)) - (g^2)'(\beta(K)) \right| = 2.55351 \cdot 10^{-15}.$$

Illustrons les trois types de points fixes de g^2 selon les valeurs de K donnés par $K = 1$, $K = 5/2$ et $K = -1$, par la figure H.8.

(b) Concluons maintenant sur la convergence de la suite (v_n) du point fixe associée à g^2 , c'est-à-dire définie par

$$v_{n+1} = g^2(v_n) \text{ et } v_0 \in \mathbb{R}. \quad (\text{H.59})$$

Considérons les intervalles de \mathbb{R} définis par

- Si $K \leq 1$,

$$I_1 =]-\infty, r(K)[, \quad (\text{H.60a})$$

$$I_2 =]r(K), +\infty[, \quad (\text{H.60b})$$

- Si $K > 1$,

$$I_3 =]-\infty, \alpha(K)[, \quad (\text{H.61a})$$

$$I_4 =]\alpha(K), r(K)[, \quad (\text{H.61b})$$

$$I_5 =]r(K), \beta(K)[, \quad (\text{H.61c})$$

$$I_6 =]\beta(K), +\infty[, \quad (\text{H.61d})$$

Remarquons d'abord que, d'après la croissance stricte de g^2 ,

$$\text{chacun des intervalles } J_k \text{ est } g^2\text{-stable.} \quad (\text{H.62})$$

En effet, par exemple, $x \in I_3$, on a $x < \alpha(K)$ et donc $g^2(x) < g^2(\alpha(K)) = \alpha(K)$. Si par exemple, $x \in I_4$, on a $\alpha(K) < x < r(K)$ et donc $g^2(\alpha(K)) < g^2(x) < g^2(r(K))$, ce qui implique $\alpha(K) < g^2(x) < r(K)$.

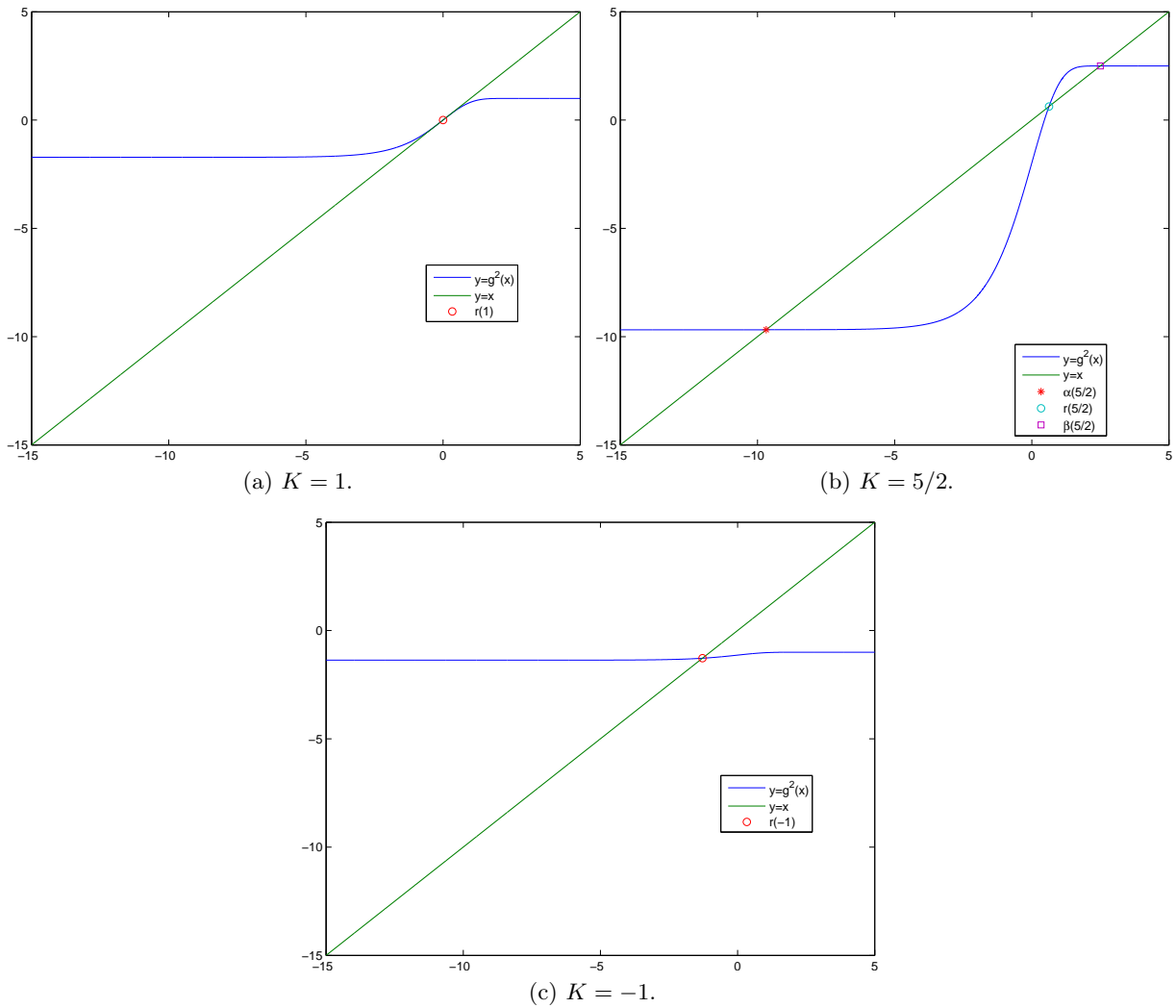


FIGURE H.8. Les graphiques de la fonction g^2 sur l'intervalle $[-15, 5]$ pour différentes valeurs de K .

Remarquons aussi que si v_0 est égal à $r(K)$ si $K \leq 1$ ou si v_0 appartient à $\{\alpha(K), r(K), \beta(K)\}$ si $K > 1$, alors la suite v_n est constante et égale à v_0 .

Supposons maintenant que v_0 appartient à l'un des intervalles I_k . Alors, pour tout n , v_n appartient à I_k , ce qui se montre par récurrence sur n en utilisant (H.62).

Enfin, v_0 appartient à l'un des intervalles I_k , alors la suite v_n est strictement monotone, à terme dans I_k . En effet, si par exemple v_0 appartient à I_4 , alors, pour tout n est dans I_4 et d'après (H.55b) appliqué à $x = v_n$, on a $v_{n+1} = g^2(v_n) < v_n$. La suite v_n étant monotone et bornée (dans $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$), elle converge vers $l \in \overline{\mathbb{R}}$. De plus, d'après (H.40) et (H.55), si v_0 appartient à I_k , pour $k \in \{1, 3, 5\}$, la suite v_n est strictement croissante et majorée dans \mathbb{R} et, pour $k \in \{2, 4, 6\}$, la suite v_n est strictement croissante et minorée dans \mathbb{R} . Dans tous les cas, la suite v_n converge donc vers l , qui est point fixe de g^2 , puisque g^2 est continue (on passe à la limite $n \rightarrow \infty$ dans (H.59)).

Or, les seuls points fixes de g^2 sont nécessairement $\alpha(K)$ ou $r(K)$ ou $\beta(K)$. La limite de v_n est donc nécessairement l'un de ces trois réels. Dans le cas où $K \leq 1$, le seul point fixe de g^2 est $r(K)$, qui est donc la valeur de l .

Si $K > 1$, montrons que l vaut $\alpha(K)$ ou $\beta(K)$.

- Si v_0 appartient à I_3 , v_n est croissante et est à valeur dans I_3 . La valeur de l est donc le seul point fixe de g^2 qui est dans l'adhérence de I_3 , c'est-à-dire $]-\infty, \alpha(K)]$, qui ne peut être que $\alpha(K)$.
- Il en est de même si v_0 appartient à I_6 avec $l = \beta(K)$.
- Si v_0 appartient à I_4 , v_n est décroissante et est à valeur dans I_4 . La valeur de l est donc le seul point fixe de g^2 qui est dans l'adhérence de I_4 , c'est-à-dire $[\alpha(K), r(K)]$. Puisque v_n est strictement décroissante, ce ne peut pas être $r(K)$ et on a donc $l = \alpha(K)$.
- Il en est de même si v_0 appartient à I_5 avec $l = \alpha(K)$.

Bref, on a montré que

$$\text{Si } K \leq 1, \text{ alors pour tout } v_0 \in \mathbb{R}, \text{ la suite } v_n \text{ converge vers } r(K). \quad (\text{H.63a})$$

$$\text{Si } K > 1 \text{ et si } v_0 = r(K), \text{ alors la suite } v_n \text{ est constante et vaut } r(K). \quad (\text{H.63b})$$

$$\text{Si } K > 1 \text{ et si } v_0 < r(K), \text{ alors la suite } v_n \text{ converge vers } \alpha(K). \quad (\text{H.63c})$$

$$\text{Si } K > 1 \text{ et si } v_0 > r(K), \text{ alors la suite } v_n \text{ converge vers } \beta(K). \quad (\text{H.63d})$$

REMARQUE H.6. On a étudié dans les remarques H.2, H.4 et H.5, les aspects attractifs de $r(K)$ et répulsifs de $\alpha(K)$ et $\beta(K)$ (des points où la dérivée (de la fonction d'itération) est en valeur absolue strictement inférieure à 1 sont dit attractifs, où la dérivée est en valeur absolue strictement supérieure à 1 sont dit répulsifs). Il est intéressant de constater que la méthode du point fixe ne converge que vers des points attractifs et non vers des point répulsifs. Attention, les aspects attractif ou répulsif ne suffisent pas à montrer la convergence ou la divergence de la méthode ; ces propriétés découlent ici de l'étude globale de la suite (voir propositions A.16 et A.17).

REMARQUE H.7. Ce résultat pouvait être démontré en utilisant la proposition F.1.

(c) Concluons par quelques simulations confirmant cela.

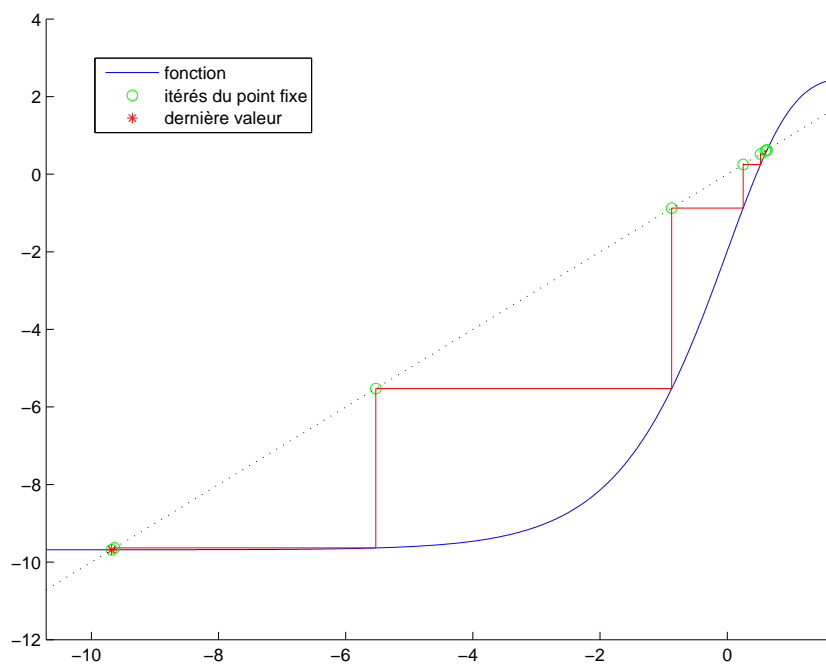


FIGURE H.9. Les 12 premières valeurs de v_n pour $K = 5/2$ et $x_0 = 5/8$.

Quelques simulations ont été faites : Voir les figures H.9, H.10 et H.11, et les tableaux H.7 H.8 et H.9.

On peut évaluer $r(-1)$ et $\alpha(5/2)$ et $\beta(5/2)$: on obtient

$$r(-1) = -1.278464542761074, \quad (\text{H.64a})$$

$$\alpha(5/2) = -9.681733635899509, \quad (\text{H.64b})$$

$$\beta(5/2) = 2.499937586791849 \quad (\text{H.64c})$$

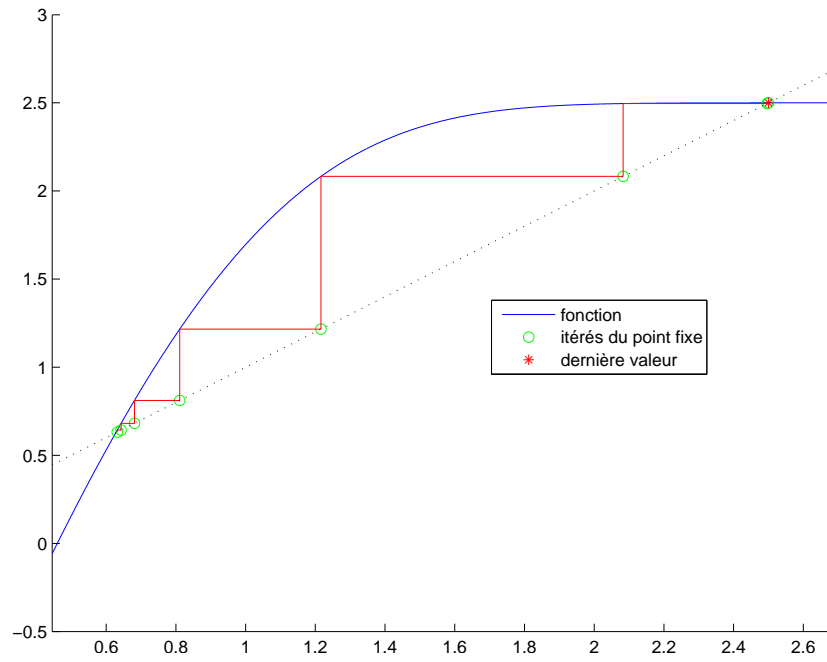


FIGURE H.10. Les 11 premières valeurs de v_n pour $K = 5/2$ et $x_0 = g(5/8)$.

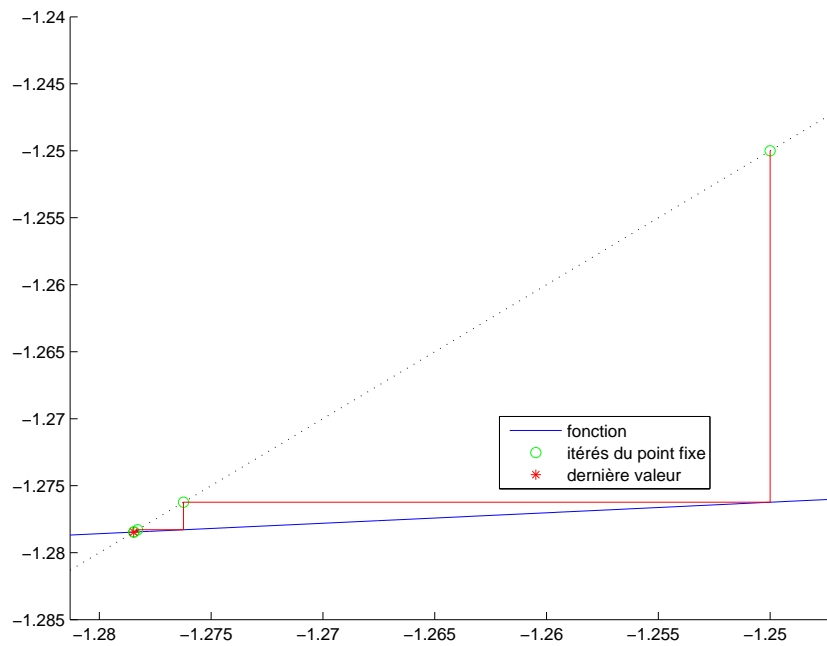


FIGURE H.11. Les 13 premières valeurs de v_n pour $K = -1$ et $x_0 = -5/4$.

et on obtient pour chacune des trois simulations :

$$|v_{13} - r(-1)| = 0, \quad (\text{H.65a})$$

$$|v_{12} - \alpha(5/2)| = 0, \quad (\text{H.65b})$$

$$|v_{11} - \beta(5/2)| = 0, \quad (\text{H.65c})$$

n	v_n
0	0.625000000000000
1	0.61909312177470
2	0.59828327950539
3	0.52413439208549
4	0.24983444603184
5	-0.87429745515325
6	-5.52727415891270
7	-9.63414264798875
8	-9.68169657765000
9	-9.68173360772356
10	-9.68173363587809
11	-9.68173363589949
12	-9.68173363589951

TABLE H.7. Les 12 premières valeurs de v_n pour $K = 5/2$ et $x_0 = 5/8$.

n	v_n
0	0.63175404256778
1	0.64275701533088
2	0.68100658482373
3	0.81100379305089
4	1.21618714121982
5	2.08284501315852
6	2.49602318550744
7	2.49993454467100
8	2.49993758447888
9	2.49993758679009
10	2.49993758679185
11	2.49993758679185

TABLE H.8. Les 11 premières valeurs de v_n pour $K = 5/2$ et $x_0 = g(5/8)$.

ce qui confirme (H.63).

(d) Étudions maintenant la convergence de la suite x_n définie par (H.24).

(i) Dans le cas où $x_0 = r(K)$, d'après (H.17c), on a pour tout n , $x_n = r(K)$.

(ii) Si $x_0 < r(K)$, d'après (H.17b) appliqué à x_0 , on a $x_1 = g(x_0) > r(K)$; puis, d'après (H.17a) appliqué à x_1 , on a $x_2 = g(x_1) < r(K)$. On montre aisément par récurrence sur n que

$$x_0 < r(K) \implies \forall n, \quad x_{2n} < r(K) \text{ et } x_{2n+1} > r(K). \quad (\text{H.66})$$

De même, on montre aisément par récurrence sur n que

$$x_0 > r(K) \implies \forall n, \quad x_{2n} > r(K) \text{ et } x_{2n+1} < r(K). \quad (\text{H.67})$$

On considère ensuite les deux suites w_n et z_n des termes de rangs pairs et impairs définies par

$$\forall n \in \mathbb{N}, w_n = x_{2n}, \quad z_n = x_{2n+1}. \quad (\text{H.68})$$

n	v_n
0	-1.250000000000000
1	-1.27623459377097
2	-1.27829148779366
3	-1.27845112280818
4	-1.27846350213932
5	-1.27846446206863
6	-1.27846453650398
7	-1.2784645427588
8	-1.27846454272345
9	-1.27846454275816
10	-1.27846454276085
11	-1.27846454276106
12	-1.27846454276107
13	-1.27846454276107

TABLE H.9. Les 13 premières valeurs de v_n pour $K = -1$ et $x_0 = -5/4$.

On alors

$$\begin{aligned}w_{n+1} &= x_{2n+2} = g(x_{2n+1}) = g(g(x_{2n})) = g^2(w_n), \\z_{n+1} &= x_{2n+3} = g(x_{2n+2}) = g(g(x_{2n+1})) = g^2(z_n),\end{aligned}$$

et autrement dit,

les deux suites w_n et z_n sont les deux suites du point fixe associées à la fonction g^2

$$\text{par } x_{n+1} = g^2(x_n) \text{ de premiers termes respectifs } x_0 \text{ et } x_1. \quad (\text{H.69})$$

Ainsi, d'après (H.66) et (H.67), on a

$$x_0 < r(K) \implies \forall n, \quad w_n < r(K) \text{ et } z_n > r(K), \quad (\text{H.70a})$$

$$x_0 > r(K) \implies \forall n, \quad w_n > r(K) \text{ et } z_n < r(K) \quad (\text{H.70b})$$

Enfin, on n'a plus qu'à utiliser (H.63) :

$$\text{Si } K \leq 1, \text{ les deux suites } w_n \text{ et } z_n \text{ convergent toutes les deux vers } r(K), \quad (\text{H.71a})$$

et donc la suite x_n converge vers $r(K)$.

$$\text{Si } K > 1, \text{ les deux suites } w_n \text{ et } z_n \text{ convergent respectivement vers } \alpha(K) \text{ et } \beta(K) \text{ si } x_0 < r(K), \quad (\text{H.71b})$$

$$\text{Si } K > 1, \text{ les deux suites } w_n \text{ et } z_n \text{ convergent respectivement vers } \beta(K) \text{ et } \alpha(K) \text{ si } x_0 > r(K), \quad (\text{H.71c})$$

et donc, $K > 1$, puisque $\alpha(K) \neq \beta(K)$, la suite x_n ne converge donc pas.

Bref, pour récapituler, on a montré que, pour tout K dans \mathbb{R} .

$$\text{Si } x_0 = r(K), \text{ la suite } x_n \text{ est constante égale à } r(K). \quad (\text{H.72a})$$

$$\text{Si } x_0 \neq r(K) \text{ et } K \leq 1, \text{ la suite } x_n \text{ converge vers } r(K). \quad (\text{H.72b})$$

$$\text{Si } x_0 \neq r(K) \text{ et } K > 1, \text{ la suite } x_n \text{ est divergente.} \quad (\text{H.72c})$$

REMARQUE H.8. Voir de nouveau la remarque H.6.

REMARQUE H.9. Ce résultat pouvait être démontré en utilisant la proposition F.2.

(e) Les simulations numériques et le corrigé des questions 2a et 2b, illustreront les résultats (H.72) établis.

◇

(a) On pose, dans cette question ;

$$a = -2, \quad (\text{H.73a})$$

$$b = -1/2, \quad (\text{H.73b})$$

$$K = -1. \quad (\text{H.73c})$$

(i) Montrons, qu'avec ces valeurs, la méthode du point fixe converge pour tout $x_0 \in [a, b]$. Il suffit d'utiliser directement la proposition A.19.

La fonction g étant décroissante, l'intervalle $[a, b]$ est g -stable si l'on a

$$a \leq g(b) \text{ et } g(a) \leq b, \quad (\text{H.74})$$

ce qu'on vérifie numériquement.

La fonction g' donnée par (H.19) étant décroissante et négative, on a donc

$$\max_{x \in [a, b]} |g'(x)| = e^{\max\{a, b\}}, \quad (\text{H.75})$$

qui est strictement plus petit que 1 si a et b sont strictement négatifs, ce qui est le cas ici. On a donc

$$k = \max_{x \in [a, b]} |g'(x)|, \quad (\text{H.76})$$

où

$$k = e^{-1/2} \text{ avec } k < 1. \quad (\text{H.77})$$

Les deux hypothèses de la proposition A.19 sont vérifiées ce qui permet de conclure.

(ii) La valeur de n telle que

$$|x_n - r| \leq \varepsilon \quad (\text{H.78})$$

avec

$$\varepsilon = 10^{-2}, \quad (\text{H.79})$$

est donnée par la proposition A.21. Avec k donné par (H.77), on obtient numériquement

$$n = 11. \quad (\text{H.80})$$

(iii)

On a affiché sur la figure H.12 page suivante, les 12 premières valeurs de x_n pour $K = -1$ et $x_0 = -5/4$. On a indiqué dans le tableau H.10 page ci-contre, les valeurs correspondantes (en allant plus loin, jusqu'à $n = 29$), en séparant les termes d'indices impairs et pairs.

Cela est conforme au résultat (H.72b), puisqu'ici, $K \leq 1$ et $r(-1)$ est donné par (H.64a). Cela est aussi conforme aux résultat (H.71a) : les deux suites de rangs pairs et impairs convergent toutes les deux vers $r(-1)$, en étant l'une inférieure, l'autre supérieure à $r(-1)$ et étant toutes les deux monotones.

REMARQUE H.10. Notons qu'il est possible de déterminer de façon explicite, grâce à la fonction W de Lambert la valeur de $r(K)$, pour tout K réel. Voir la section D.2. On a ici

$$a = 1, \quad (\text{H.81a})$$

$$b = 1, \quad (\text{H.81b})$$

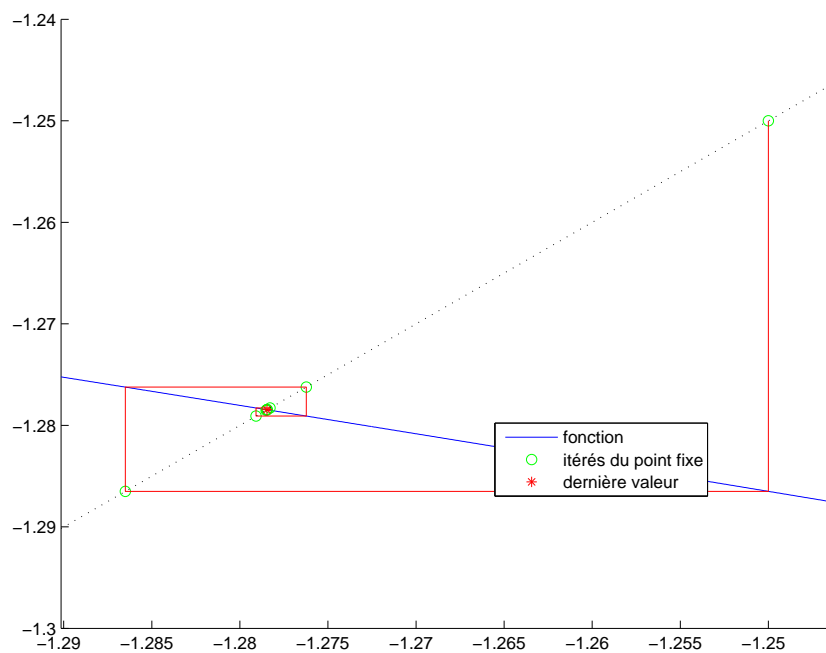
$$c = -K \quad (\text{H.81c})$$

et donc

$$\Delta = e^K \in \mathbb{R}_+^*, \quad (\text{H.82})$$

et on a donc d'après (D.11), la solution unique $r(K)$, donnée par

$$r(K) = K - W_0(e^K). \quad (\text{H.83})$$

FIGURE H.12. Les 12 premières valeurs de x_n pour $K = -1$ et $x_0 = -5/4$.

n	x_{2n}	x_{2n+1}
0	-1.25000000000000	-1.28650479686019
1	-1.27623459377097	-1.27908619735840
2	-1.27829148779366	-1.27851273660342
3	-1.27845112280818	-1.27846827976719
4	-1.27846350213932	-1.27846483253749
5	-1.27846446206863	-1.27846456523106
6	-1.27846453650398	-1.27846454450345
7	-1.27846454227588	-1.27846454289618
8	-1.27846454272345	-1.27846454277155
9	-1.27846454275816	-1.27846454276189
10	-1.27846454276085	-1.27846454276114
11	-1.27846454276106	-1.27846454276108
12	-1.27846454276107	-1.27846454276107
13	-1.27846454276107	-1.27846454276107
14	-1.27846454276107	-1.27846454276107

TABLE H.10. Les 30 premières valeurs de x_n pour $x_0 = -5/4$ et $K = -1$.

Cela est confirmé par matlab :

```
solve('e^x=-x+K','x')
```

qui donne

$$\frac{-\text{LambertW}(\ln(e) e^{\ln(e)K}) + \ln(e)K}{\ln(e)}$$

la fonction `lambertw` est programmée sous matlab et on peut donc utiliser directement (H.83). Numériquement, on a

$$r(-1) = -1.2784645427610739, \quad (\text{H.84})$$

ce qui confirme la valeur finale du tableau H.10. De plus, pour n donné par (H.80), on a

$$|x_n - r(-1)| = 2.24699 \cdot 10^{-8},$$

ce qui confirme *a posteriori* le choix de n défini par la majoration (H.78).

(b) On pose, dans cette question ;

$$a = 1/4, \quad (\text{H.85a})$$

$$b = 1, \quad (\text{H.85b})$$

$$K = 5/2. \quad (\text{H.85c})$$

(i) Comme dans la question 2a, la fonction g' donnée par (H.19) est décroissante et négative et on a donc

$$\min_{x \in [a,b]} |g'(x)| = e^{\min\{a,b\}}, \quad (\text{H.86})$$

qui est strictement plus grand que 1 si l'un des réels a et b est strictement positif, ce qui est le cas ici. On a donc

$$k = \min_{x \in [a,b]} |g'(x)|, \quad (\text{H.87})$$

où

$$k = e^{1/4} \text{ avec } k > 1. \quad (\text{H.88})$$

Par ailleurs, on a

$$\text{signe}(f(a)f(b)) = -1, \quad (\text{H.89})$$

et donc l'intervalle $[a, b]$ contient l'unique racine $r(K)$ de f . Donc, de (H.87) et (H.88), on peut déduire que

$$|g'(r(K))| > 1 \quad (\text{H.90})$$

et le point $r(K)$ est dit répulsif. *Attention*, l'égalité (H.90) ne permet pas *a priori* d'affirmer que la méthode du point fixe est divergente.

Pour cela, on peut utiliser par exemple la proposition A.17. Ses trois hypothèses sont vérifiées.

En effet, on pose $I = [a, b]$ et on successivement :

— g est définie sur \mathbb{R} donc sur I .

— La propriété (A.30b) est vérifiée. En effet, cette propriété est équivalente à

$$g(\mathbb{R} \setminus I) \subset \mathbb{R} \setminus I. \quad (\text{H.91})$$

La fonction g étant décroissante, l'intervalle $\mathbb{R} \setminus [a, b]$ est g -stable si l'on a

$$g(b) \leq a \text{ et } g(a) \geq b, \quad (\text{H.92})$$

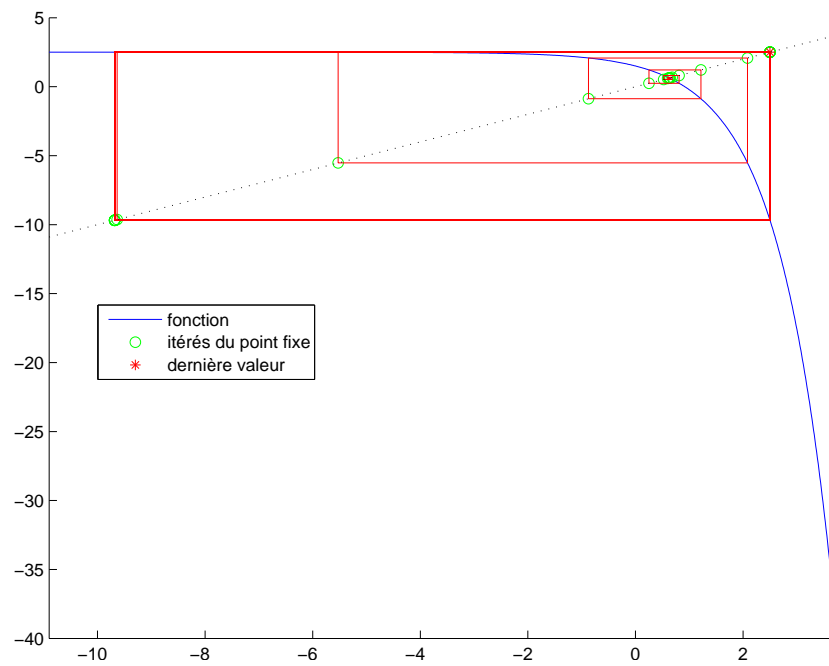
ce qu'on vérifie numériquement.

— Enfin, la propriété (A.30c) est vérifiée d'après (H.87) et (H.88).

On peut donc conclure que la suite x_n diverge pour tout $x_0 \in [a, b] \setminus \{r(K)\}$.

La preuve de ce résultat a été établie dans le cas général pour tout $K > 1$ et pour tout $x_0 \neq r(K)$. Voir (H.72c).

◇

FIGURE H.13. Les 30 premières valeurs de x_n pour $x_0 = 5/8$ et $K = 5/2$.

n	x_{2n}	x_{2n+1}
0	0.625000000000000	0.63175404256778
1	0.61909312177470	0.64275701533088
2	0.59828327950539	0.68100658482373
3	0.52413439208549	0.81100379305089
4	0.24983444603184	1.21618714121982
5	-0.87429745515325	2.08284501315852
6	-5.52727415891270	2.49602318550744
7	-9.63414264798875	2.49993454467100
8	-9.68169657765000	2.49993758447888
9	-9.68173360772356	2.49993758679009
10	-9.68173363587809	2.49993758679185
11	-9.68173363589949	2.49993758679185
12	-9.68173363589951	2.49993758679185
13	-9.68173363589951	2.49993758679185
14	-9.68173363589951	2.49993758679185

TABLE H.11. Les 30 premières valeurs de x_n pour $x_0 = 5/8$ et $K = 5/2$.

(ii)

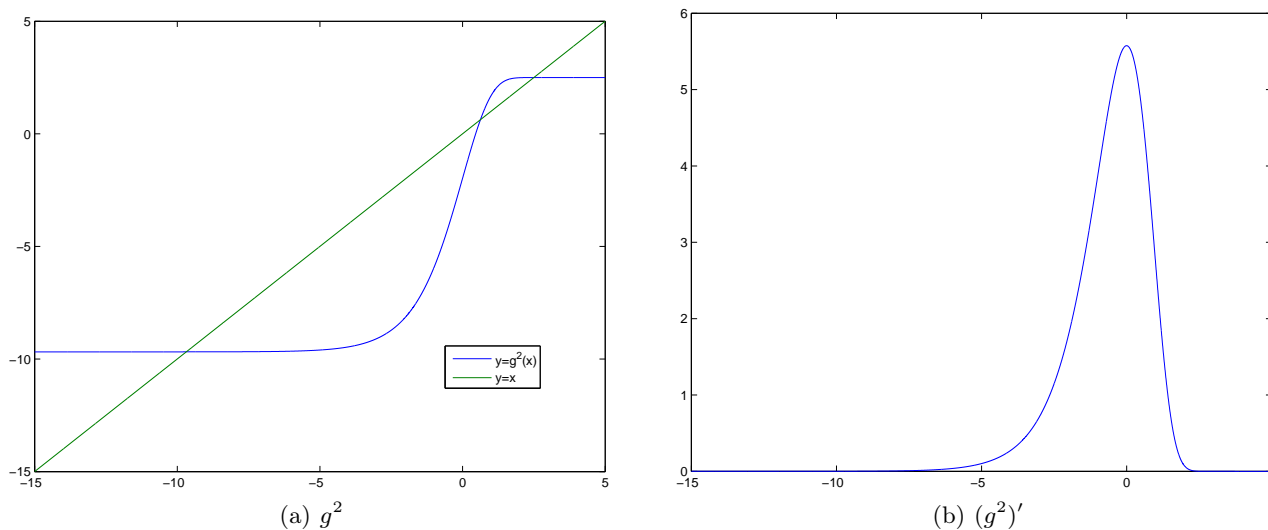


FIGURE H.14. Les graphiques des fonction g^2 et $(g^2)'$ sur l'intervalle $[-15, 5]$.

On a affiché sur la figure H.13 page précédente, les 30 premières valeurs calculées avec les paramètres donnés par (H.85) et x_0 donné par

$$x_0 = 5/8. \quad (\text{H.93})$$

On a indiqué dans le tableau H.11 page précédente, les valeurs correspondantes, en séparant les termes d'indices impairs et pairs.

(A)

On constate sur la figure H.13 page précédente et surtout sur le tableau H.11 page précédente, que la suite x_n semble ne pas converger. Plus précisément, il semble apparaître que la suite des termes de rangs pairs convergerait en décroissant vers la valeur

$$l_p = -9.6817336358995085, \quad (\text{H.94})$$

et que la suite des termes de rangs impairs convergerait en croissant vers la valeur

$$l_i = 2.4999375867918490, \quad (\text{H.95})$$

(B)

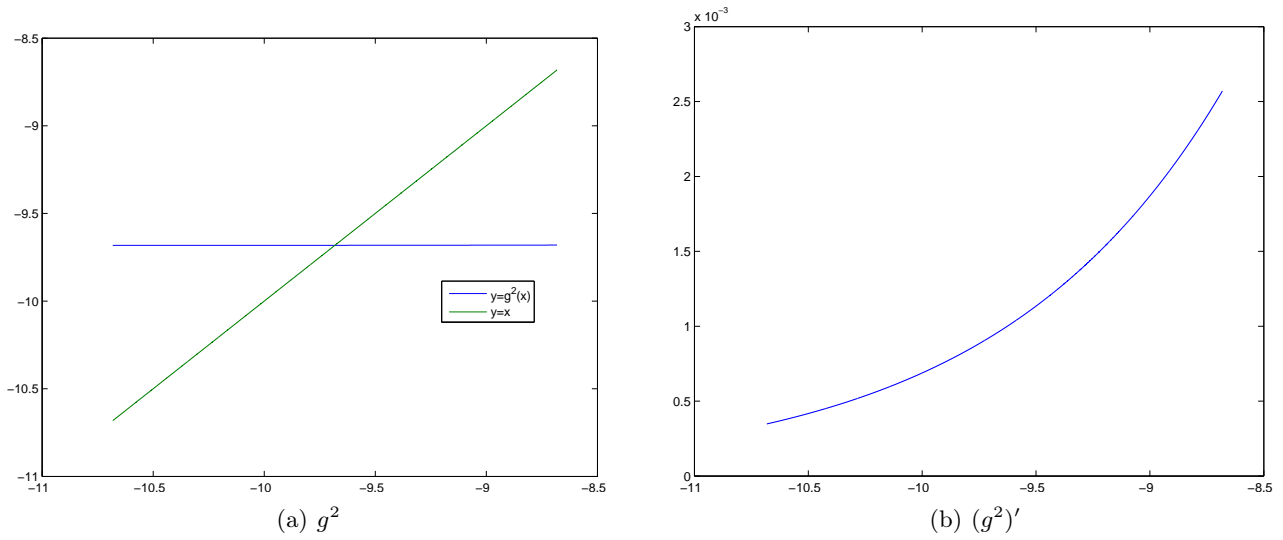
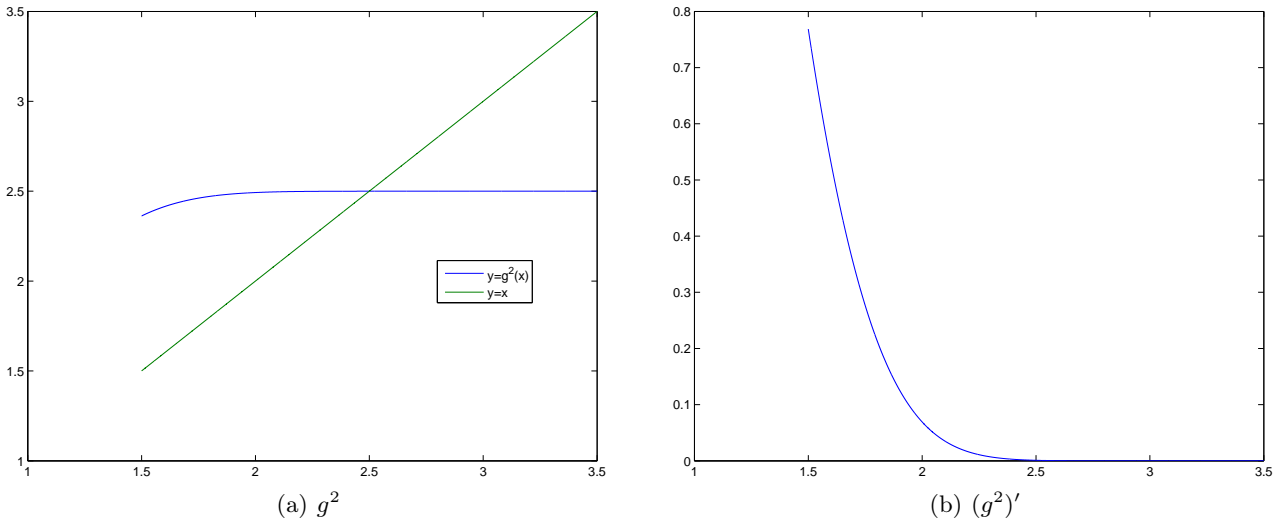
Tâchons d'expliquer de façon qualitative les observations faites dans la question 2(b)iiA grâce aux figures H.14, H.15 et H.16. En fait, la suite x_{2n} des termes de rangs pairs vérifie la relation de récurrence

$$x_{2n+2} = g(x_{2n+1}) = g(g(x_{2n})) = g^2(x_{2n}),$$

où g^2 est donnée par (H.25). De même, la suite x_{2n+1} des termes de rangs impairs vérifie la relation de récurrence

$$x_{2n+3} = g(x_{2n+2}) = g(g(x_{2n+1})) = g^2(x_{2n+1}).$$

Le comportement de la fonction g^2 permet donc de prédire la convergence des suites x_{2n} et x_{2n+1} . On constate sur la figure H.14, que la fonction g^2 semble avoir trois points fixes dont les valeurs numériques sont proches des trois valeurs données par (H.64). Il semblerait que le plus grand et le plus petit point fixe soient des points attractifs pour la fonction g^2 comme le montrent les figures H.15 et H.16, et que le point fixe intermédiaire soit répulsif

FIGURE H.15. Les graphiques des fonction g^2 et $(g^2)'$ sur l'intervalle $[-10.68173, -8.68173]$.FIGURE H.16. Les graphiques des fonction g^2 et $(g^2)'$ sur l'intervalle $[1.49994, 3.49994]$.

comme le montre la figure (14(b)). On peut confirmer cela puisque que l'on peut vérifier numériquement qu'en ces trois points fixes, on a

$$g'(r(5/2)) = 2.567854985668169, \quad (\text{H.96a})$$

$$g'(\alpha(5/2)) = 0.000760301077059, \quad (\text{H.96b})$$

$$g'(\beta(5/2)) = 0.000760301077059. \quad (\text{H.96c})$$

On pourrait aussi montrer que les hypothèses de la proposition A.19 sont valables pour la fonction g^2 sur des intervalles autour des valeurs du premier et du dernier point fixe de g^2 données par (H.64b) et (H.64c). Ainsi, il y aurait bien convergence des suites x_{2n} et x_{2n+1} vers ces valeurs.

De façon plus rigoureuse, il suffit d'utiliser les résultats (H.71b) et (H.72c), qui prévoient la divergence de la suite x_n et la convergence de x_{2n} et x_{2n+1} vers $\alpha(K)$ et $\beta(K)$ (distincts), puisque $x_0 < r(K)$. Enfin, on confirme cela de façon numérique puisque pour $n = 15$, on a

$$\begin{aligned} |x_{2n} - l_p| &= 0, \\ |x_{2n+1} - l_i| &= 0. \end{aligned}$$

- (iii) Pour résoudre (H.9) dont on sait que la solution $r(5/2)$ existe et est unique, on ne peut donc utiliser la méthode du point fixe qui diverge. En revanche, on peut utiliser la méthode de la dichotomie sur l'intervalle $[a, b]$, qui converge ici vers $r(5/2)$ d'après (H.89). On choisit n pour que la méthode de dichotomie fournisse $r(5/2)$ avec une erreur inférieure à ε donné par

$$\varepsilon = 10^{-16},$$

en utilisant la proposition A.14. On obtient pour le n -ième milieu

$$x_n = 0.627352959583406,$$

et, en utilisant (H.83), qui fournit,

$$r(5/2) = 0.627352959583406,$$

on a

$$|x_n - r(5/2)| = 0.$$

Étude et calcul de l tel que $l = \cos l$ sous la forme d'un problème corrigé

Énoncé

(1) On considère la suite $(u_n)_{n \in \mathbb{N}}$ définie par $u_0 \in \mathbb{R}$ et, pour tout $n \geq 0$, $u_{n+1} = \cos(u_n)$.

(a) Montrer que l'intervalle $I = [\cos(1), 1]$ est stable par la fonction g donnée par $g(x) = \cos(x)$. Par la suite, on pourra donc supposer (quitte à poser $u_0 = u_2$) que u_2 appartient à I .

(b) Étudier la convergence de la suite $(u_n)_{n \in \mathbb{N}}$.

(c) (i) On pose

$$\varepsilon_1 = 10^{-1} \text{ et } \varepsilon_2 = 10^{-15}. \quad (\text{I.1})$$

Notons α la limite de la suite. Déterminer deux entiers n_1 et n_2 tels que

$$|u_{n_1} - \alpha| \leq \varepsilon_1 \text{ et } |u_{n_2} - \alpha| \leq \varepsilon_2. \quad (\text{I.2})$$

(ii) Pour $u_0 = 10$, calculer numériquement les $n_1 + 1$ premières valeurs de u_n . Déterminer l'écart entre u_{n_1} et α , solution de $x = \cos(x)$ déterminé de façon exacte donné par

$$\alpha = 0.7390851332151606 \quad (\text{I.3})$$

et conclure.

(2) On veut maintenant déterminer de façon numérique la valeur de α donnée par (I.3) mais plus rapidement que dans la question 1.

(a) Quelle méthode itérative¹ pourriez-vous utiliser pour répondre à cette question ?

(b) On définit la fonction h par

$$\forall x \in \mathbb{R}, \quad h(x) = x - \cos(x). \quad (\text{I.4})$$

On *admet* que h admet un unique zéro sur \mathbb{R} . On note (w_n) la suite associée à cette méthode. Précisez la fonction G telle que

$$\forall x \in \mathbb{R}, \quad w_{n+1} = G(w_n). \quad (\text{I.5})$$

(c) (i) Pourquoi la méthode étudiée est-elle quadratique au voisinage de l'unique zéro de h sur l'intervalle I , défini par

$$I = [0, \pi/2] ? \quad (\text{I.6})$$

(ii) Démontrer que l'inégalité

$$\forall n \in \mathbb{N}, \quad |w_{n+1} - l| \leq D |w_n - l|^2 \quad (\text{I.7})$$

est satisfaite pour

$$D = \frac{1}{2} \max_{x \in I} |G''(x)|, \quad (\text{I.8})$$

1. Il n'y a pas d'ambiguïté sur cette question, car on n'a vu qu'une seule méthode répondant à cela !

en admettant que la méthode étudiée converge pour tout x_0 appartenant à l'intervalle I vers l'unique zéro de h .

(iii) En fait, l'intervalle I est un peu trop grand. On se place désormais sur l'intervalle J défini par

$$J = [\nu, \pi/2 - \nu], \text{ où } \nu = 0.3000, \quad (\text{I.9})$$

sur lequel tout ce qui a été fait précédemment est encore valable. On fournit alors la majoration suivante de la constante D , valable sur l'intervalle J :

$$D \leq 0.73741536. \quad (\text{I.10})$$

On considère les nombres ε_1 et ε_2 définis par (I.1). Déterminer deux entiers m_1 et m_2 tels que

$$|u_{m_1} - \alpha| \leq \varepsilon_1 \text{ et } |u_{m_2} - \alpha| \leq \varepsilon_2, \quad (\text{I.11})$$

et comparer avec les résultats de la question 1(c)i.

Commentez !

(iv) Déterminez les $m_1 + 1$ premières valeurs de la suite w_n pour $w_0 = 0.3000$ et concluez grâce à la valeur de α donnée par (I.3).

Corrigé

(1) (a) Définissons la fonction g par

$$\forall x \in \mathbb{R}, \quad g(x) = \cos(x). \quad (\text{I.12})$$

Posons

$$I = [\cos(1), 1] \quad (\text{I.13})$$

qui est g -stable. Voir la définition A.18 du cours. En effet, $\cos(1) \approx 0.5403023$ appartient à $[0, \pi/2]$ et 1 appartient à $[0, \pi/2]$ et puisque \cos est décroissant sur $[0, \pi/2]$, pour tout $x \in [\cos(1), 1]$, on a $\cos(x) \in [\cos(1), \cos(\cos(1))] \approx [\cos(1), 0.8575532] \subset [\cos(1), 1]$.

REMARQUE I.1. On peut aussi montrer cela sans calculer les valeurs approchées. En effet, $1 \in]0, \pi/2[$ et donc $\cos(1) \in]0, 1[$. Ainsi, $[\cos(1), 1] \subset [0, 1] \subset [0, \pi/2]$ et par décroissance de \cos sur $[0, \pi/2]$, on a $g([\cos(1), 1]) = [\cos(1), \cos(\cos(1))] \subset [\cos(1), 1]$ puisque $\cos(\cos(1)) \leq 1$ et $\cos(\cos(1)) \neq 1$, sinon $\cos(1) = 0$.

Si $u_0 \in \mathbb{R}$, on a $u_1 = \cos(u_0) \in [-1, 1]$ et u_2 appartient donc à $\cos([-1, 1]) = \cos([0, 1]) = [\cos(1), \cos(0)] = I$. Donc,

$$n \geq 2, \quad u_n \in I, \quad (\text{I.14})$$

ce que l'on montre par récurrence, puisque I est g -stable. Si la suite u_n converge, la fonction g étant continue, c'est nécessairement vers un point fixe de g . Définissons la fonction h par

$$\forall x \in \mathbb{R}, \quad h(x) = x - \cos(x). \quad (\text{I.15})$$

Puisque $h(\cos(1))h(1) \approx -0.1458395 < 0$, la fonction h étant continue, cela implique que h admet au moins une racine sur $[\cos(1), 1]$ et donc g admet au moins un point fixe, noté α et on vérifie que

$$\alpha \in I \quad (\text{I.16})$$

On montrera plus bas que α est unique.

REMARQUE I.2. Un autre choix consiste à choisir

$$I = [1/2, 1]. \quad (\text{I.17})$$

Si $u_0 \in \mathbb{R}$, on a $u_1 = \cos(u_0) \in [-1, 1]$ et u_2 appartient donc à $\cos([-1, 1]) = \cos([0, 1]) = [\cos(1), \cos(0)]$, qui est inclus dans I . En effet, on a $0 < 1 < \pi/3 < \pi/2$ et donc $\cos(1) > \cos(\pi/3) = 1/2$. De plus, I est g -stable. En effet, si x est dans $[1/2, 1]$, on a $\cos(x)$ dans $[\cos(1), \cos(1/2)]$ inclus dans $[\cos(\pi/3), \cos(0)] = I$. Ainsi, (I.14) reste vraie. Enfin, (I.16) reste vraie puisque $h(1/2)h(1) \approx -0.1735738 < 0$.

REMARQUE I.3. Un autre choix consiste à choisir

$$I = [0, 1]. \quad (\text{I.18})$$

Si $u_0 \in \mathbb{R}$, on a $u_1 = \cos(u_0) \in [-1, 1]$. Si $u_1 \in [-1, 0[$, on peut remplacer u_1 par $-u_1 \in [0, 1]$ puisque $u_2 = \cos(u_1) = \cos(-u_1)$. On peut donc supposer que $u_1 \in [0, 1]$. On a donc I est g -stable. En effet, on a $0 < 1 < \pi/2$ et donc $\cos([0, 1]) = [\cos(1), \cos(0)] = [\cos(1), 1] \subset [0, 1]$. Ainsi, (I.14) reste vraie.

Par la suite, on pourra donc supposer (quitte à appliquer la règle (I.22)) que u_2 (ou que u_1 dans le cas de la remarque I.3) appartient à I .

(b) On peut donc maintenant conclure sur la convergence de la suite $(u_n)_{n \in \mathbb{N}}$.

On a, pour tout $x \in I$, $|g'(x)| = |\sin(x)|$. La fonction \sin est croissante sur $I = [\cos(1), 1]$ et est comprise, sur cet intervalle, entre $\sin(\cos(1)) \approx 0.5143953$ et $\sin(1) \approx 0.8414710$. L'inégalité de l'hypothèse (A.31b) du cours est donc vraie avec

$$k = \sin(1). \quad (\text{I.19})$$

Puisque $0 < 1 < \pi/2$, on a $\sin(1) \in]0, 1[$. On laisse au lecteur le soin de vérifier que ce raisonnement est encore valable si l'on choisit I donné dans la remarque I.2 ou I.3, puisque

$$0 \leq \sin(0) \leq \sin(1/2) \leq \sin(1).$$

Il suffit donc d'appliquer la proposition A.19 du cours qui implique à la fois la convergence de la suite et l'unicité du point fixe α de g .

REMARQUE I.4. Le choix de I donné par (I.17) ou (I.18) serait aussi valable.

REMARQUE I.5. On aurait pu se passer de l'utilisation de la proposition A.19 du cours et montrer la convergence de la suite à la main, en raisonnant comme suit :

(i)

Étudions donc tout d'abord la fonction g , définie par (I.12). Si on définit la fonction h par (I.15), alors, on a

$$\forall x \in \mathbb{R}, \quad h'(x) = 1 + \sin(x), \quad (\text{I.20})$$

qui est positive sur \mathbb{R} et nulle pour $x = \pi/2 + 2k\pi$ où k entier. De plus $h(\pm\infty) = \pm\infty$. Ainsi h est strictement croissante sur \mathbb{R} et n'admet qu'un seul zéro, noté α et qui est donc l'unique point fixe de g .

(ii)

Voir par exemple les figures I.1 et I.2

On a, pour tout $x \in I$,

$$x > \alpha \implies g(x) < \alpha, \quad (\text{I.21a})$$

$$x < \alpha \implies g(x) > \alpha. \quad (\text{I.21b})$$

En effet, si $x \in]\alpha, 1[\subset [0, \pi/2]$, puisque $g = \cos$ est décroissante sur $[0, \pi/2]$, on a $g(x) \in]\cos(1), \cos \alpha]$ et donc $g(x) < \alpha$. Il en est de même si $x < \alpha$. On peut supposer sans perte de généralité que $u_0 \in I$, quitte à appliquer la règle suivante :

$$\text{on remplace } u_0 \text{ par } u_2. \quad (\text{I.22})$$

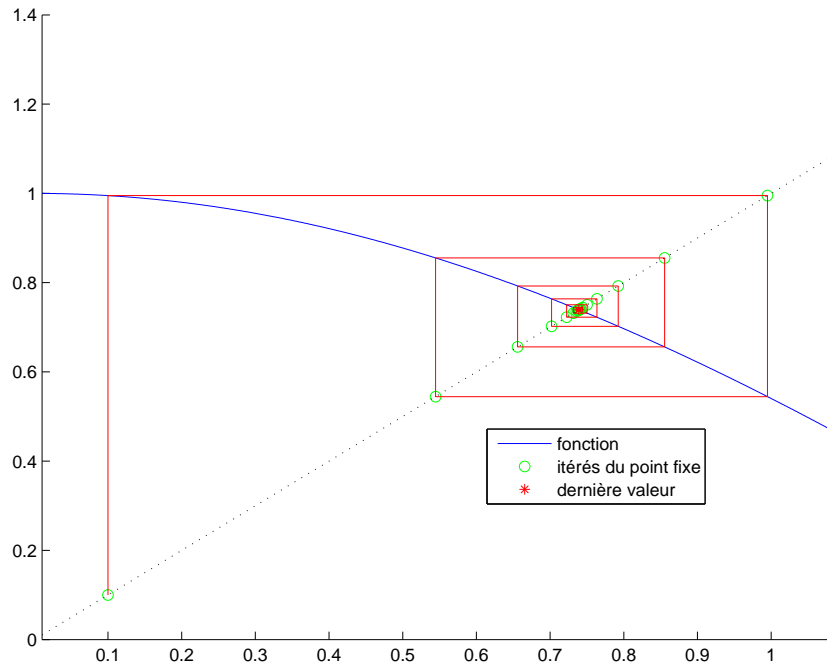


FIGURE I.1. Le tracé graphique des valeurs de u_n pour $u_0 = 0.100$.

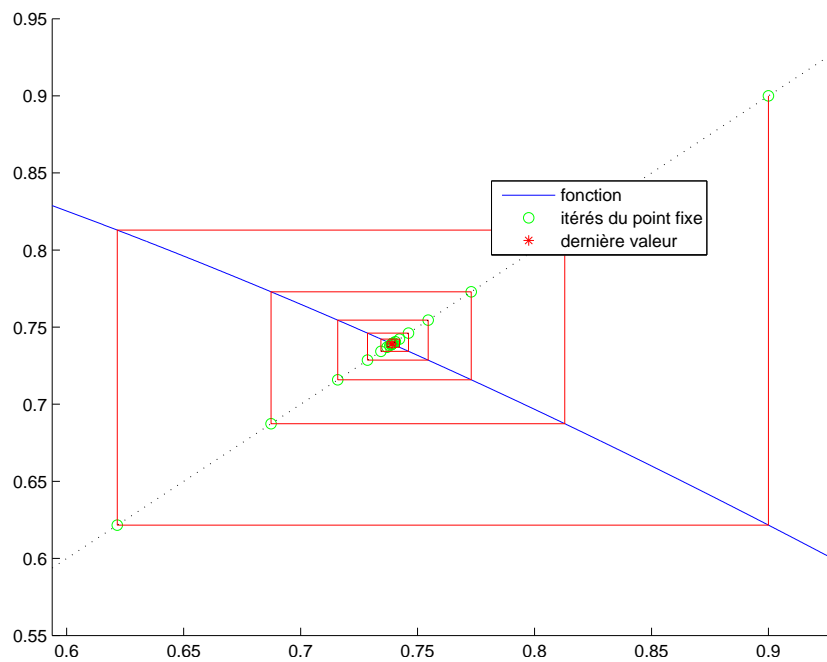


FIGURE I.2. Le tracé graphique des valeurs de u_n pour $u_0 = 0.900$.

Ainsi, si $u_0 = \alpha$, alors la suite u_n est constante et égale à α . Sinon, si $u_0 < \alpha$, alors d'après (I.21), $u_1 = g(u_0) > \alpha$ et $u_2 < \alpha$ et par récurrence

$$\forall n \in \mathbb{N}, \quad u_{2n} \in [\cos(1), \alpha[, \quad u_{2n+1} \in]\alpha, 1]. \tag{I.23}$$

De même, on montrer que, si $u_0 > \alpha$,

$$\forall n \in \mathbb{N}, \quad u_{2n+1} \in [\cos(1), \alpha[, \quad u_{2n} \in]\alpha, 1]. \quad (\text{I.24})$$

Enfin, si on pose

$$f(x) = g(g(x)) - x, \quad (\text{I.25})$$

on a

$$f'(x) = g'(g(x))g'(x) - 1$$

et donc

$$f'(x) = \sin(\cos(x)) \sin(x) - 1. \quad (\text{I.26})$$

Pour $x \in I$, $\cos(x) \in I$, $\sin(x) \in [\sin(\cos(1)), \sin(1)]$ et $\sin(\cos(x)) \in [\sin(\cos(1)), \sin(1)] \approx [0.5143953, 0.8414710]$. On a donc $\sin(\cos(x)) \sin(x) < \sin^2(1)$ et $f'(x) \leq \sin^2(1) - 1 \approx -0.2919266 < 0$ et donc f est strictement décroissante sur I . Or $f(\alpha) = g(g(\alpha)) - \alpha = 0$, donc pour tout $x \in I$

$$x > \alpha \implies g(g(x)) < x, \quad (\text{I.27a})$$

$$x < \alpha \implies g(g(x)) > x. \quad (\text{I.27b})$$

REMARQUE I.6. Comme précédemment, on peut montrer cela sans les valeurs numériques. En effet, si $x \in [\cos(1), 1]$, on a $\cos(x) \in [\cos(1), 1] \subset]0, 1[$ et donc $\sin(\cos(x)) \in]0, 1[$. Il en est de même pour $\sin(x)$. Ainsi, en particulier $0 < \sin(\cos(x)) \sin(x) < 1$ et $f'(x) < 0$.

REMARQUE I.7. Ce raisonnement est encore valable pour le choix (I.17) ou (I.18).

Soit $n \in \mathbb{N}$. Si $u_{2n} < \alpha$, alors, d'après (I.27), on a $u_{2n+2} > u_{2n}$. En combinant cela, (I.23) et (I.24), on montre donc par récurrence que si $u_0 < \alpha$,

$$\forall n \in \mathbb{N}, \quad u_{2n} \in [\cos(1), \alpha[\text{ et la suite } u_{2n} \text{ est croissante;} \quad (\text{I.28a})$$

$$\forall n \in \mathbb{N}, \quad u_{2n+1} \in]\alpha, 1] \text{ et la suite } u_{2n+1} \text{ est décroissante;} \quad (\text{I.28b})$$

De même, si $u_0 > \alpha$,

$$\forall n \in \mathbb{N}, \quad u_{2n+1} \in [\cos(1), \alpha[\text{ et la suite } u_{2n+1} \text{ est croissante;} \quad (\text{I.29a})$$

$$\forall n \in \mathbb{N}, \quad u_{2n} \in]\alpha, 1] \text{ et la suite } u_{2n} \text{ est décroissante;} \quad (\text{I.29b})$$

Dans tous les cas, les deux suites u_{2n} et u_{2n+1} sont monotones, bornées (car dans I) et donc convergentes, chacune vers un zéro de f , puisque à la limite $f(l) = g(g(l)) - l = 0$. Puisque f est strictement décroissante, f n'admet au plus qu'un seul zéro; or α est un zéro de f donc les deux suites u_{2n} et u_{2n+1} convergent toutes les deux vers α et u_n aussi.

REMARQUE I.8. Ce résultat pouvait être démontré en utilisant la proposition F.2.

◇

(c) (i) On pose

$$\varepsilon_1 = 10^{-1} \text{ et } \varepsilon_2 = 10^{-15}. \quad (\text{I.30})$$

On considère la valeur de k donnée par (I.19) et a et b définis par

$$a = \cos(1) \text{ et } b = 1. \quad (\text{I.31})$$

Utilisons la proposition A.21 du cours. Ainsi, les deux entiers n_1 et n_2 tels que $|u_n - l| \leq \varepsilon_1$ et $|u_n - l| \leq \varepsilon_2$ sont donnés respectivement par

$$n_1 = 9 \text{ et } n_2 = 196. \quad (\text{I.32})$$

Attention, puisque que u_0 , n'est pas dans l'intervalle $I = [a, b]$ et puisque l'on a appliqué la règle (I.22), il faut rajouter 2 à chacun des entiers définis précédemment de sorte que

$$n_1 = 11 \text{ et } n_2 = 198. \quad (\text{I.33})$$

n	u_n
0	10.000000000000000
1	-0.839071529076452
2	0.668153917531387
3	0.784966720933852
4	0.707411791257438
5	0.760046415906649
6	0.724804033046625
7	0.748629360225378
8	0.732622462482921
9	0.743423001354606
10	0.736156148510940
11	0.741054959438722

TABLE I.1. 12 premières valeurs de la méthode du point fixe pour $g(x) = \cos(x)$.

(ii)

Donnons dans le tableau (I.1), les $n_1 + 1$ premières valeurs de u_n .

Rappelons que α , unique solution de $x = \cos(x)$ déterminé de façon exacte², est donné par

$$\alpha = 0.7390851332151606$$

On a

$$|\alpha - u_{n_1}| \approx 1.9698 \cdot 10^{-3},$$

ce qui est bien inférieur à la valeur de $\varepsilon_1 = 10^{-1}$ donnée dans l'énoncé.

On pourrait remarquer que pour $n = 198$, on a

$$|\alpha - u_n| \approx 3.9540 \cdot 10^{-16},$$

ce qui est strictement inférieur à la valeur de $\varepsilon_2 = 10^{-15}$ donnée dans l'énoncé.

(2) (a) La seule méthode connue d'ordre strictement plus grand que 1 est, dans ce cours, la méthode de Newton, censée être quadratique, donc convergeant plus rapidement que la méthode du point fixe précédemment et qui n'est que linéaire.

(b) On rappelle que la fonction h est définie par (I.15). La valeur de α définie dans la question (1) est donc un zéro de la fonction h qui est unique sur \mathbb{R} d'après l'étude faite dans la remarque I.5 page 127 et qui comme précédemment est noté α . Pour l'intervalle I définie par

$$I = [0, \pi/2], \tag{I.34}$$

on a

$$\text{signe}(h(0)h(\pi/2)) = -1, \tag{I.35}$$

et l'unique zéro de h est donc dans l'intérieur de l'intervalle I . La relation liant w_{n+1} à w_n est donnée par la définition (A.53) appliquée à la fonction h . On a donc ici d'après (I.15) et (I.20),

$$x - \frac{h(x)}{h'(x)} = x - \frac{x - \cos(x)}{1 + \sin(x)} = \frac{x + x \sin(x) - x + \cos(x)}{1 + \sin(x)},$$

2. en fait très précise grâce à l'une des fonction `solve` ou `fzero` de matlab.

et donc

$$w_{n+1} = G(w_n), \quad (\text{I.36a})$$

où

$$G(x) = \frac{x \sin(x) + \cos(x)}{1 + \sin(x)}. \quad (\text{I.36b})$$

(c) Cette correction est très proche de la la correction de l'exercice 3.4 page 38 de TD et notamment de la question 2 page 39 auxquelles on renvoie pour plus de détail.

(i) Voir la proposition A.30. On y a vu que la la méthode de Newton est quadratique ssi $h'(\alpha) \neq 0$ et si $h''(\alpha) \neq 0$. On sait que $h'(\alpha)$ est non nul, (d'après la question 1 puisque $h' > 0$), et le cours assure que $G'(\alpha) = 0$. Ainsi, la méthode est au moins quadratique.

Rappelons que ce résultat est fondé sur le calcul suivant : On a, là où $h' \neq 0$,

$$G' = 1 - \left(\frac{h}{h'}\right)' = 1 - \frac{h'^2 - h''h}{h'^2} = \frac{1}{h'^2} (h'^2 - h'^2 + h''h),$$

et donc

$$G' = \frac{h''h}{h'^2},$$

et en particulier, puisque $h'(\alpha) \neq 0$,

$$G'(\alpha) = \frac{h''(\alpha)h(\alpha)}{h'^2(\alpha)}.$$

Ainsi, $G'(\alpha) = 0$ si et seulement si $h(\alpha) = 0$, ce qui est vrai.

De plus, elle est exactement quadratique si et seulement si $G''(\alpha) \neq 0$, ce qui est équivalent à $h''(\alpha) \neq 0$.

Rappelons que ce résultat est fondé sur le calcul suivant : On peut montrer que

$$G'' = \frac{1}{(h')^4} \left((h')^3 h'' + hh''' (h')^2 - hh'' \left((h')^2 \right)' \right),$$

En particulier, en α , puisque $h'(\alpha)$ est non nul et que $h(\alpha)$ est nul, on a

$$G''(\alpha) = \frac{1}{(h'(\alpha))^4} \left((h'(\alpha))^3 h''(\alpha) \right) = \frac{h''(\alpha)}{h'(\alpha)},$$

et donc

$$G''(\alpha) = \frac{h''(\alpha)}{h'(\alpha)}.$$

Ainsi, $G''(\alpha) \neq 0$ si et seulement si $h''(\alpha) \neq 0$.

D'après (I.20), h' est on nul à l'intérieur de l'intervalle I auquel on sait que la racine α appartient.

Ainsi, d'après la proposition A.30, la méthode de Newton est au moins quadratique. Enfin, d'après (I.20), on a

$$\forall x \in \mathbb{R}, \quad h''(x) = \cos(x), \quad (\text{I.37})$$

qui est non nul à l'intérieur de l'intervalle I à laquelle on sait que la racine α appartient.

Ainsi, d'après la proposition A.30, la méthode de Newton est exactement quadratique.

(ii) Voir les propositions A.26 et A.30.

On y a vu que le développement de Taylor de la fonction g sur $[x^*, x_n]$ permet de montrer que

$$|x_{n+1} - x^*| \leq D |x_n - x^*|^2, \quad (\text{I.38})$$

avec

$$D = \frac{1}{2} \max_{x \in I} |G''(x)| \quad (\text{I.39})$$

en admettant d'abord *a priori* que la suite x_n converge et que les x_n sont dans l'intervalle I .

Rappelons que ce résultat est fondé sur le calcul suivant : On applique la formule de Taylor-Lagrange à G avec $a = \alpha$, $b = w_n$ à l'ordre $p = 1$: il existe $\xi_n \in]\alpha, w_n[$ tel que

$$G(w_n) = G(\alpha) + G'(\alpha)(w_n - \alpha) + \frac{1}{2}G''(\xi_n)(w_n - \alpha)^2,$$

et on donc

$$w_{n+1} - \alpha = G'(\alpha)(w_n - \alpha) + \frac{1}{2}G''(\xi_n)(w_n - \alpha)^2,$$

et donc, puisque $G'(\alpha) = 0$

$$w_{n+1} - \alpha = \frac{1}{2}G''(\xi_n)(w_n - \alpha)^2.$$

Si la suite w_n converge, alors, puisque $\xi_n \in [r, w_n]$, $\xi_n \in [r, w_n]$ tend vers r et puisque G'' est continue $G''(\xi_n)$ tend vers $G''(r)$ et donc

$$\lim_{n \rightarrow +\infty} \frac{|e_{n+1}|}{|e_n|^2} = \frac{1}{2}G''(r).$$

et donc

$$\exists D, \quad \forall n \in \mathbb{N}, \quad \frac{|e_{n+1}|}{|e_n|^2} \leq D.$$

REMARQUE I.9. Justifions maintenant rigoureusement la convergence de la suite w_n pour tout x_0 appartenant à I . Il suffit pour cela d'invoquer la première proposition de l'annexe intitulée "Convergence globale de la méthode de Newton" de [bastienmmmatcours], appliquée à la fonction h , dont les différentes hypothèses sont vérifiées sur l'intervalle I . En effet, :

- (A) l'hypothèse 1) est clairement vérifiée ;
- (B) l'hypothèse 2) est clairement vérifiée d'après (I.35) ;
- (C) l'hypothèse 3) est vérifiée d'après (I.20), puisque le sinus est positif ou nul sur I ;
- (D) l'hypothèse 4) est vérifiée d'après (I.37). Attention, en $\pi/2$, le cosinus s'annule mais on peut montrer que si on part de $x_0 = \pi/2$ ou de $x_0 < \pi/2$, on reste dans tous les cas dans un intervalle strictement inclus dans $[0, \pi/2[$.
- (E) l'hypothèse 5) est vérifiée d'après les valeurs numériques suivantes :

$$\left(\frac{\pi}{2}\right)^{-1} \frac{|h(0)|}{|h'(0)|} \approx 0.63661977236758 < 1,$$

$$\left(\frac{\pi}{2}\right)^{-1} \frac{|h(\frac{\pi}{2})|}{|h'(\frac{\pi}{2})|} \approx 0.50000000000000 < 1.$$

◇

- (iii) Il suffit de raisonner comme dans la question 4 page 41 de la correction de l'exercice de TD 3.4 ou dans la proposition A.24.

On a

$$D = \max_{x \in I} \frac{|\cos(x)|}{|1 + \sin(x)|}.$$

On a donc, pour tout $x \in I = [0, \pi/2]$:

$$\frac{|\cos(x)|}{|1 + \sin(x)|} = \frac{\cos(x)}{1 + \sin(x)} \leq \frac{\cos(x)}{1} \leq \cos x \leq 1.$$

Dans ce cas, il est nécessaire pour appliquer la proposition A.24 que la majoration (A.42) soit vérifiée ; or, ici, en prenant $e_0 = \pi/2$ et $p = 2$ (méthode quadratique), on a

$$|e_0|D^{\left(\frac{1}{p-1}\right)} = \frac{\pi}{2},$$

qui n'est pas majoré strictement par 1. Si au contraire, on se place sur l'intervalle $J = [\nu, \pi/2 - \nu]$ avec $0 < \nu < \pi/4$, on a clairement pour tout $x \in J$:

$$\frac{|\cos(x)|}{|1 + \sin(x)|} = \frac{\cos(x)}{1 + \sin(x)}$$

et donc

$$D \leq \frac{\cos(\nu)}{1 + \sin \nu}. \quad (\text{I.40})$$

Avec le choix de ν donné dans l'énoncé, on vérifie d'une part que la première proposition de l'annexe intitulée "Convergence globale de la méthode de Newton" de **[bastienmmatercours]**, s'applique pour la fonction h , comme fait dans la remarque I.9, et d'autre part que l'on a bien, d'après (I.40), le résultat donné dans l'énoncé :

$$D \leq 0.73741536. \quad (\text{I.41})$$

◇

Dans ce cas, la majoration (A.42) est vérifiée puisque numériquement en prenant $p = 2$ (méthode quadratique), on a

$$\left(\frac{\pi}{2} - 2\nu\right) D^{\left(\frac{1}{p-1}\right)} = 0.715880114974; \quad (\text{I.42})$$

et on peut donc appliquer la proposition A.24. Il ne reste plus qu'à appliquer l'équation (A.44) appelée ici :

$$n = \left\lceil \frac{1}{\ln p} \ln \left(\frac{\ln \frac{\varepsilon}{\gamma}}{\ln \delta} \right) \right\rceil, \quad (\text{I.43})$$

avec γ et δ définis par l'équation (A.40) c'est-à-dire

$$\gamma = C^{\left(\frac{1}{1-p}\right)}, \quad (\text{I.44a})$$

$$\delta = |e_0| C^{\left(\frac{1}{p-1}\right)}, \quad (\text{I.44b})$$

Numériquement, on a pour $p = 2$, on a pour $|e_0| = \frac{\pi}{2} - 2\nu$, on a

$$\gamma = 0.737415351928,$$

$$\delta = 0.715880114974.$$

puis pour les deux entiers m_1 et m_2 , associé à ε_1 et ε_2 , on a

$$m_1 = 3,$$

$$m_2 = 7,$$

qui sont des valeurs beaucoup plus faibles que celles données par (I.32)! Cela nous montre *a posteriori* l'efficacité accrue de la méthode de Newton!

(iv) Concluons par des simulations numériques.

n	w_n
0	0.3000000000000000
1	0.805848141739496
2	0.740002131837112
3	0.739085318708854
4	0.739085133215168
5	0.739085133215161

TABLE I.2. 6 premières valeurs de la méthode de Newton.

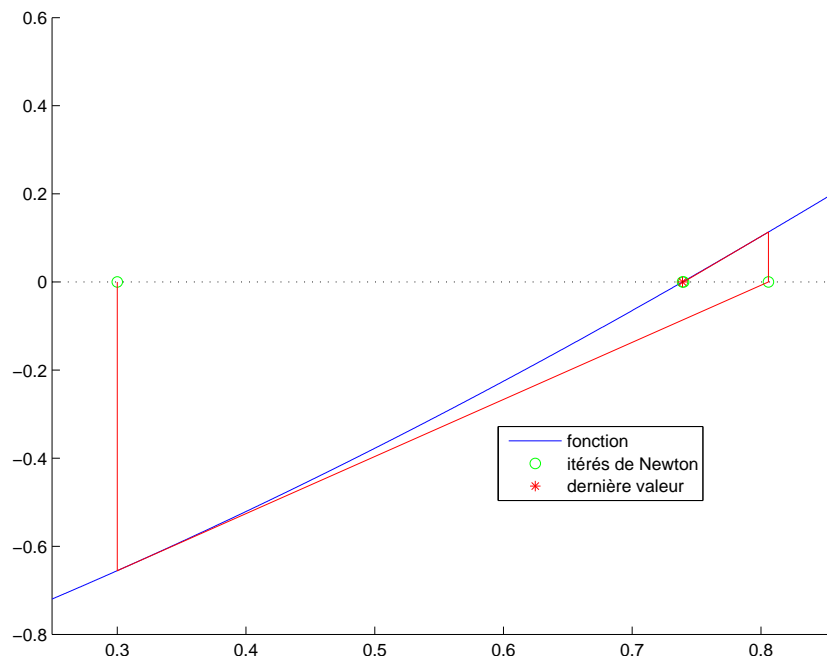


FIGURE I.3. Le tracé graphique des valeurs de w_n de la méthode de Newton pour $w_0 = 0.3000$.

Voir la figure I.3 et le tableau I.2 page précédente

Enfin, on a

$$|\alpha - u_{m_1}| \approx 9.1699 \cdot 10^{-4},$$

ce qui est bien inférieur à la valeur de $\varepsilon_1 = 10^{-1}$ donnée dans l'énoncé.

On pourrait remarquer que pour $n = 5$, on a

$$|\alpha - u_n| \approx 3.0638 \cdot 10^{-17},$$

ce qui est strictement inférieur à la valeur de $\varepsilon_2 = 10^{-15}$ donnée dans l'énoncé.

Dégénérescence de la méthode de Newton et méthode de Newton modifiée

Étudions ce qu'il se passe quand $f'(r) = 0$, dans la méthode de Newton (voir définition A.29 et propriété A.30).

Les résultats de cette annexe utilisent les notions de racines multiples de fonctions, qui généralisent celles de polynômes, le tout étant rappelé dans l'annexe K page 141.

PROPOSITION J.1. *Soit $m \in \mathbb{N}^*$. Si f est de classe \mathcal{C}^{m+2} sur un intervalle du type $[r - \varepsilon, r + \varepsilon]$, si r est un zéro d'ordre¹ m alors,*

(1) *Il existe h de classe \mathcal{C}^2 sur $[r - \varepsilon, r + \varepsilon]$ telle que*

$$\forall x \in [r - \varepsilon, r + \varepsilon], \quad f(x) = (x - r)^m h(x), \quad (\text{J.1a})$$

avec

$$h(r) \neq 0. \quad (\text{J.1b})$$

(2) *Quitte à supposer ε assez petit,*

$$f' \text{ est non nulle sur } [r - \varepsilon, r + \varepsilon] \setminus \{r\}. \quad (\text{J.2})$$

(3) (a) *La méthode de Newton est exactement linéaire.*

(b) *Si, de plus, f est de classe \mathcal{C}^{m+3} sur $[r - \varepsilon, r + \varepsilon]$, alors, la méthode de Newton modifiée définie par $x_{n+1} = g(x_n)$ où*

$$\forall x \in [r - \varepsilon, r + \varepsilon], \quad g(x) = \begin{cases} x - m \frac{f(x)}{f'(x)}, & \text{si } x \neq r, \\ r, & \text{si } x = r. \end{cases} \quad (\text{J.3})$$

est au moins quadratique. Si

$$h'(r) \neq 0, \quad (\text{J.4})$$

elle est exactement quadratique

On pourra aussi consulter

<https://math-linux.com/mathematiques/resolution-numerique-des-equations-non-lineaires/article/methode-de-newton>

DÉMONSTRATION DE LA PROPOSITION J.1.

(1) Les résultats du point 1 proviennent de la section K.2 de l'annexe K, notamment le lemme K.3 avec $p = m + 2$.

(2) On déduit de (J.1a)

$$\forall x \in [r - \varepsilon, r + \varepsilon], \quad f'(x) = m(x - r)^{m-1}h(x) + (x - r)^m h'(x). \quad (\text{J.5})$$

De (J.1b), on déduit donc (J.2).

1. Voir annexe K et particulièrement la section K.2 et la définition K.6.

(3) Considérons la fonction d'itération ϕ_A définie par, pour tout $A \in \mathbb{R}^*$:

$$\forall x \in [r - \varepsilon, r + \varepsilon] \setminus \{r\}, \quad \phi_A(x) = x - A \frac{f(x)}{f'(x)} \quad (\text{J.6})$$

On a alors, grâce à (J.1a)

$$\forall x \in [r - \varepsilon, r + \varepsilon] \setminus \{r\}, \quad \phi_A(x) = x - A \frac{(x-r)^m h(x)}{m(x-r)^{m-1} h(x) + (x-r)^m h'(x)}, \quad (\text{J.7})$$

et donc

$$\forall x \in [r - \varepsilon, r + \varepsilon] \setminus \{r\}, \quad \phi_A(x) = x - A \frac{(x-r)h(x)}{mh(x) + (x-r)h'(x)}, \quad (\text{J.8})$$

et donc, par prolongement par continuité

$$\forall x \in [r - \varepsilon, r + \varepsilon], \quad \phi_A(x) = x - A \frac{h(x)(x-r)}{mh(x) + (x-r)h'(x)}. \quad (\text{J.9})$$

D'après (J.1b), ϕ_A est de classe \mathcal{C}^1 puisque son dénominateur est non nul en r . Ainsi,

$$\phi'_A(x) = 1 - A \frac{(h(x) + (x-r)h'(x))(mh(x) + (x-r)h'(x)) - (h(x)(x-r))(mh'(x) + h'(x) + (x-r)h''(x))}{(mh(x) + (x-r)h'(x))^2} \quad (\text{J.10})$$

D'après (J.9)

$$\phi_A(x) = r \quad (\text{J.11})$$

et d'après (J.10)

$$\phi'_A(x) = r - \frac{Amh^2(r)}{m^2h^2(r)}$$

et donc

$$\phi'_A(x) = 1 - \frac{A}{m}. \quad (\text{J.12})$$

On peut montrer sous matlab que

$$\phi_A(r) = r - A \times 0 = r, \quad (\text{J.13a})$$

$$\phi'_A(r) = 1 - Am^{-1}. \quad (\text{J.13b})$$

On retrouve donc les résultats (J.11) et (J.12). De plus, si f est de classe \mathcal{C}^{m+3} , alors h est de classe \mathcal{C}^3 et ϕ_A est de classe \mathcal{C}^2 . On peut montrer grâce à matlab que

$$\forall x \in [r - \varepsilon, r + \varepsilon] \setminus \{r\}, \quad \phi''_A(x) = A \frac{N(x)}{D(x)},$$

avec

$$D(x) = \left(mh(x) + \left(\frac{d}{dx} h(x) \right) x - \left(\frac{d}{dx} h(x) \right) r \right)^3,$$

et $N(x)$ une expression, complexe et non affichée, dépendant de r , x , m et des trois premières dérivées de h et que

$$\phi''_A(r) = \frac{2Ah'(r)}{m^2h(r)}. \quad (\text{J.14})$$

REMARQUE J.2. Informatiquement, on ne peut poser $\phi_A(r) = r$, car r est, en principe, inconnu ! De, l'expression théorique donnée par (J.8) n'est pas accessible. Seule, l'expression donnée par (A.53) est accessible. L'usage de cette expression peut aussi, compte tenu des arrondis de calculs, donner des résultats moins précis que ce que prévoit la théorie. D'après (J.2), cette expression est légitime. Cependant, si $x_{n+1} = x_n$ et si on se trouve dans le cas où $x_{n+1} = x_n$, alors on a

$$A \frac{f(x_n)}{f'(x_n)} = 0$$

et donc $f(x_n) = 0$ et on s'arrête là !

On peut conclure et démontrer maintenant les résultats des cas 3a et 3b.

- (a) La méthode de Newton classique correspond à $A = 1$ et $\phi_A = g$, définie par (A.53). g est de classe \mathcal{C}^1 . On a, d'après (J.12)

$$|g'(r)| = 1 - \frac{1}{m},$$

et on a $m > 1$ donc $1/m < 1$ et $1 - 1/m > 0$. On a aussi $1 - 1/m < 1$. Ainsi

$$|g'(r)| \in]0, 1[.$$

D'après la proposition A.25, la méthode de Newton est donc exactement linéaire.

- (b) La méthode de Newton modifiée correspond à $A = m$ et $\phi_A = g$, définie par (J.3). Si f est de classe \mathcal{C}^{m+3} , g est de classe \mathcal{C}^3 . On alors, grâce à (J.12), $g'(r) = 1 - m/m = 0$ et d'après la proposition A.26, cette méthode de Newton modifiée est au moins quadratique. De plus, dans ce cas, grâce à (J.14), on a

$$g''(r) = \frac{2h'(r)}{mh(r)}, \quad (\text{J.15})$$

ce qui permet de conclure, grâce à (J.4).

□

On pourra consulter les deux exemples ci-dessous.

EXEMPLE J.3. Considérons la fonction f définie par

$$f(x) = (x - \alpha)^p(x - \beta), \quad (\text{J.16})$$

où β et α sont deux réels distincts et $p \in \mathbb{N}^*$. Si $p \geq 2$, il est clair que α est un zéro de multiplicité $p \geq 2$ de f .

On choisit les valeurs suivantes :

$$\alpha = 2,$$

$$\beta = 1,$$

$$p = 2.$$

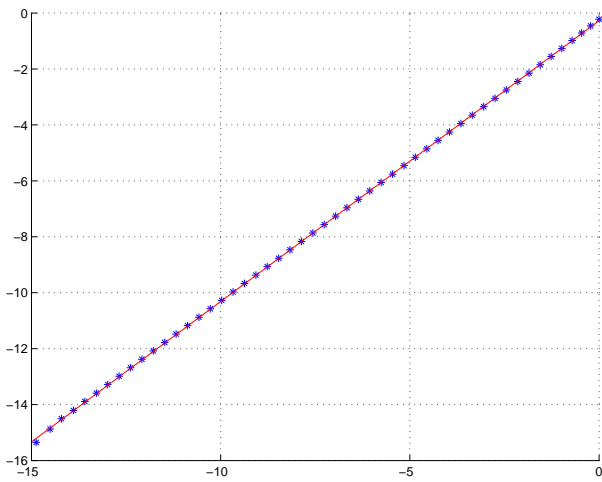
Pour g définie par (J.3), on obtient sous Matlab

$$\begin{aligned} \lim_{x \rightarrow \alpha} g(x) &= 2, \\ \lim_{x \rightarrow \alpha} g'(x) &= 0, \\ \lim_{x \rightarrow \alpha} g''(x) &= 1, \end{aligned}$$

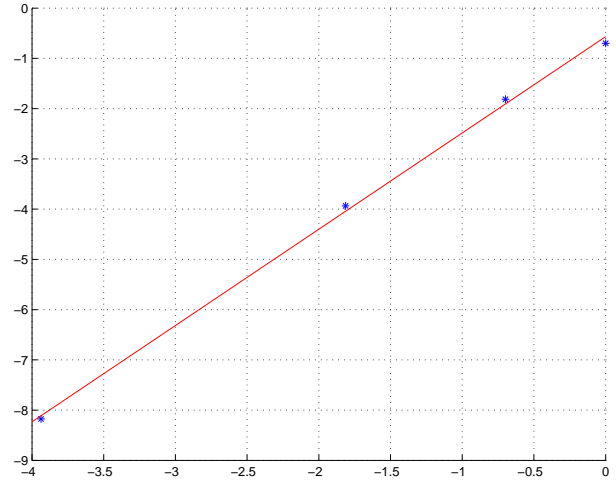
ce qui confirme que l'ordre de la méthode de Newton modifiée est exactement 2.

On étudie f au voisinage de α . On obtient les itérés de la méthode de Newton dans le tableau J.1 et ceux de la méthode modifiée dans le tableau J.2, qui semble mieux se comporter. Afin de comparer le comportement des deux méthodes, on procède exactement comme dans le corrigé de l'exercice de TD 3.5. On obtient alors les deux nuages $\ln - \ln$ des points représentés sur les figures J.1. Enfin, les deux mesures de pentes de ces nuages fournissent les corrélations 0.99998221 et 0.99937031 ainsi que les deux pentes 1.00377621 et 1.91625254 ce qui confirme l'ordre 1 de la méthode de Newton et l'ordre 2 de la méthode modifiée.

EXEMPLE J.4. On pourra consulter l'exercice 3.5 des TD où la méthode de Newton modifiée est utilisée et comparée à la méthode de Newton habituelle dans la remarque 3.6 du corrigé.



(a) Méthode de Newton habituelle



(b) Méthode de Newton modifiée

FIGURE J.1. Le nuage de points $\ln - \ln$.

n	x_n
0	3.0000000000000000
1	2.6000000000000000
2	2.347368421052632
3	2.193516663631397
4	2.104014284855781
5	2.054346842021410
6	2.027856158851746
7	2.014114290149136
8	2.007105915824401
9	2.003565448288779
10	2.001785885343129
11	2.000893737887931
12	2.000447068368469
13	2.000223584118280
14	2.00011180452414
15	2.000055905400748
16	2.000027953481662
17	2.000013976936172
18	2.000006988516924
19	2.000003494270672
20	2.000001747138388
21	2.000000873569957
22	2.000000436785169
23	2.000000218392632
24	2.000000109196328
25	2.000000054598167
26	2.000000027299084
27	2.000000013649542
28	2.000000006824771
29	2.000000003412386
30	2.000000001706193
31	2.000000000853097
32	2.000000000426549
44	2.000000000000104
45	2.000000000000052
46	2.000000000000026
47	2.000000000000013
48	2.000000000000007
49	2.000000000000004
50	2.000000000000002
51	2.000000000000001
52	2.000000000000000
53	2.000000000000000

TABLE J.1. Itérations de la méthode de Newton

n	x_n
0	3.000000000000000
1	2.200000000000000
2	2.015384615384615
3	2.000115673799884
4	2.000000006689053
5	2.000000000000000

TABLE J.2. Itérations de la méthode de Newton modifiée

Racines multiples

Dans cette annexe, nous rappelons tout d'abord la notion de racines multiples usuelles pour un polynôme puis nous l'étendons à une fonction quelconque

K.1. Racines multiples d'un polynôme

On pourra consulter par exemple https://fr.wikipedia.org/wiki/Racine_d%27un_polyn%C3%B4me ou [RDO93, section 6.4.3 2)].

On a la définition-proposition suivante

DÉFINITION K.1 (Ordre de multiplicité, racine simple, racine multiple). Soit P un polynôme, à coefficients réels ou complexes et a un nombre réel ou complexe. Alors,

- (1) le plus grand entier $m \in \mathbb{N}^*$ tel que P soit divisible par $(X - a)^m$ est appelé l'ordre, ou la multiplicité, de la racine a relativement à P ;
- (2) cet entier m est caractérisé par l'existence d'un polynôme Q tel que $P = (X - a)^m Q$ et $Q(a) \neq 0$;
- (3) cet entier m est caractérisé par

$$P(a) = P'(a) = \dots = P^{(m-1)}(a) = 0, \quad (\text{K.1a})$$

$$P^{(m)}(a) \neq 0. \quad (\text{K.1b})$$

- (4) on dit que a est racine simple de P si $m = 1$ et racine multiple si $m > 1$.

DÉMONSTRATION DE L'ÉQUIVALENCE DES CAS 1,2 ET 3.

- (1) Remarquons que l'énoncé du cas 1 est équivalent à

$$P \text{ est divisible par } (X - a)^m \text{ et } P \text{ n'est pas divisible par } (X - a)^{m+1}. \quad (\text{K.2})$$

Démontrons donc l'équivalence du cas 2 et de (K.2), ce qui est fait dans [RDO93, section 6.4.3 2) Théorème II et définition], que l'on rappelle ici :

- (a) Démontrons que le cas 2 implique (K.2).

Il est clair, dans ce cas, que P est divisible par $(X - a)^m$

Montrons maintenant que P n'est pas divisible par $(X - a)^{m+1}$. Faisons la division Euclidienne de Q par $X - a$: il existe donc un polynôme S et un réel (ou un complexe) λ tel

$$Q = (X - a)S + \lambda.$$

Si on choisit $X = a$, on a $\lambda = Q(a)$ et donc

$$Q = (X - a)S + Q(a), \quad (\text{K.3})$$

dont on déduit que

$$Q(X - a)^m = (X - a)^{m+1}S + Q(a)(X - a)^m.$$

Du cas 2, on déduit donc que

$$P = (X - a)^m Q = (X - a)^{m+1}S + Q(a)(X - a)^m$$

et donc que le reste de la division de P par $(X - a)^{m+1}$ est le polynôme $Q(a)(X - a)^m$ qui est non nul puisque $Q(a) \neq 0$. Ainsi, P n'est pas divisible par $(X - a)^{m+1}$.

(b) Montrons maintenant que (K.2) implique le cas 2. L'hypothèse se traduit par $P = Q(X - a)^m$ avec Q non divisible par $X - a$. Or, d'après (K.3) dire que Q est divisible par $X - a$ revient à dire que $Q(a) = 0$. Ainsi, $Q(a) \neq a$.

(2) Démontrons que le cas 2 est équivalent au cas 3, ce qui est fait dans [RDO93, section 6.4.3 2) Théorème III], que l'on rappelle ici :

Rappelons la formule de Taylor à l'ordre p appliquée au polynôme P (de degré p). Voir par exemple [DB22, Formule de Taylor-Lagrange dans le chapitre intitulé "Équations non-linéaires"]. Elle donne ici : il existe ξ tel que

$$P = \sum_{n=0}^p \frac{P^{(n)}(a)}{n!} (X - a)^n + \frac{P^{(p+1)}(\xi)}{(p+1)!} (X - a)^{p+1}.$$

et donc, puisque $P^{(p+1)} = 0$:

$$\begin{aligned} P &= \sum_{n=0}^p \frac{P^{(n)}(a)}{n!} (X - a)^n, \\ &= \sum_{n=0}^{m-1} \frac{P^{(n)}(a)}{n!} (X - a)^n + \frac{P^{(m)}(a)}{m!} (X - a)^m + \sum_{n=m+1}^p \frac{P^{(n)}(a)}{n!} (X - a)^n, \end{aligned}$$

et donc

$$P = \sum_{n=0}^{m-1} \frac{P^{(n)}(a)}{n!} (X - a)^n + (X - a)^m \left(\frac{P^{(m)}(a)}{m!} + \sum_{n=m+1}^p \frac{P^{(n)}(a)}{n!} (X - a)^{n-m} \right)$$

Cette expression met en valeur le quotient

$$Q = \frac{P^{(m)}(a)}{m!} + \sum_{n=m+1}^p \frac{P^{(n)}(a)}{n!} (X - a)^{n-m},$$

et le reste

$$R = \sum_{n=0}^{m-1} \frac{P^{(n)}(a)}{n!} (X - a)^n$$

de la division de P par $(X - a)^m$. En particulier, il vient

$$Q(a) = \frac{P^{(m)}(a)}{m!}$$

D'après le cas 2, a est racine d'ordre m ssi $P = (X - a)^m Q$ et $Q(a) \neq 0$, cela est aussi équivalent à $R = 0$ et $Q(a) \neq 0$ et donc en comparant avec l'expression du reste R et de $Q(a)$ à (K.1).

□

On déduit de la définition (K.1), le lemme suivant, laissé à la sagacité du lecteur :

LEMME K.2. *Le nombre a est racine simple de P ssi $P(a) = 0$ et $P'(a) \neq 0$.*

K.2. Racines multiples d'une fonction quelconque

Soient $a \in \mathbb{R}$, $\varepsilon > 0$ et On considère désormais une fonction f définie sur un intervalle $I_\varepsilon = [a - \varepsilon, a + \varepsilon]$. On donne alors le résultat suivant :

LEMME K.3. *Soient deux entiers $m, p \in \mathbb{N}^*$ tels que $m \leq p$ et f une fonction de I_ε dans \mathbb{R} , de classe \mathcal{C}^p , vérifiant*

$$f(a) = f'(a) = \dots = f^{(m-1)}(a) = 0. \quad (\text{K.4})$$

Alors la fonction g définie par

$$\forall t \in I_\varepsilon, \quad g(t) = \begin{cases} \frac{f(t)}{(t-a)^m} & \text{si } t \neq a, \\ \frac{f^{(m)}(a)}{m!} & \text{si } t = a, \end{cases} \quad (\text{K.5})$$

est de classe \mathcal{C}^{p-m} .

DÉMONSTRATION. Voir [RDO88, section 6.7.2 6)]. □

On en déduit le résultat suivant :

LEMME K.4. *On suppose que f vérifie de plus*

$$f^{(m)}(a) \neq 0, \quad (\text{K.6})$$

alors,

$$\forall t \in I_\varepsilon, \quad f(t) = (t-a)^m g(t), \quad (\text{K.7})$$

avec

$$g(a) \neq 0. \quad (\text{K.8})$$

DÉMONSTRATION. Il suffit de remarquer que $g(a) = \frac{f^{(m)}(a)}{m!} \neq 0$. □

Enfin, on a le lemme suivant :

LEMME K.5. *Soient un entier $m \in \mathbb{N}^*$, f et g deux fonctions de I_ε dans \mathbb{R} , de classe \mathcal{C}^p , vérifiant (K.7) et (K.8). Alors, f vérifie (K.4) et (K.6).*

DÉMONSTRATION. Soit $k \leq m - 1$. Pour dériver k fois (K.7) il suffit d'appliquer la formule de Leibniz :

$$(uv)^{(k)} = \sum_{l=0}^k \binom{l}{k} u^{(l)} v^{(k-l)}. \quad (\text{K.9})$$

Cette formule, tout à fait identique à la formule du binôme de Newton, est démontrée par exemple dans https://fr.wikipedia.org/wiki/Formule_de_Leibniz et https://fr.wikipedia.org/wiki/R%C3%A8gle_du_produit. On applique cette formule à (K.7) avec $u = (t-a)^m$ et $v = g$. Chacune des dérivées l -ième pour $l \leq k \leq m - 1$ de u est donc nulle en a et on a donc

$$f^{(k)}(a) = 0,$$

et on a donc (K.4). Si on prend $k = m$, alors toutes les dérivées en a sont nulles sauf la m -ième dérivée (qui correspond à $l = m$ dans (K.9)) qui vaut $m!$. On a donc

$$f^{(m)}(a) = m!g(a),$$

qui est non nul d'après (K.8). □

Autrement dit, les résultats du lemmes K.3, K.4 et K.5 généralisent l'équivalence des cas 2 et 3 de la définition K.1. Ces lemmes correspondent donc au cas où f présente une racine dite d'ordre m en a .

On peut donc donner la définition suivante d'une racine d'ordre m d'une fonction f .

DÉFINITION K.6. Soient f une fonction suffisamment régulière sur I_ε . Les deux assertions suivantes sont équivalentes

- (1) Il existe une fonction g suffisamment régulière sur I_ε vérifiant (K.7) et (K.8).
- (2) On a (K.4) et (K.6).

Dans ce cas, on dit que a est racine simple de f si $m = 1$ et racine multiple si $m > 1$.

Un théorème de point fixe

Rappelons par exemple le théorème [BM03, Théorème 4.8] :

THÉORÈME L.1 (condition suffisante de convergence de la méthode du point fixe).

Voir désormais le théorème A.19, fondé sur l'hypothèse il existe un réel k de $[0, 1[$ tel que

$$\forall x \in [\alpha, \beta] \quad |g'(x)| \leq k. \quad (\text{L.1})$$

DÉMONSTRATION. Voir preuve dans [BM03]. □

De cette preuve, il ressort que, pour tout $n \in \mathbb{N}$, on a

$$|x_{n+1} - l| \leq k |x_n - l| \quad (\text{L.2})$$

dont on déduit par récurrence sur n que pour tout $n \in \mathbb{N}$, on a

$$|x_n - l| \leq k^n |x_0 - l|, \quad (\text{L.3})$$

ce qui fournit aussi

$$|x_n - l| \leq k^n |\beta - \alpha|, \quad (\text{L.4})$$

et donc de retrouver le fait que x_n tend vers l . De plus, cette inégalité permet de déterminer le rang n à partir duquel

$$|x_n - l| \leq \varepsilon, \quad (\text{L.5})$$

pour tout $\varepsilon > 0$.

Dans le cas où l'intervalle $[\alpha, \beta]$ est trop grand, l'estimation (L.3) n'est pas utile et on lui préfère la propriété suivante :

THÉORÈME L.2 (Seconde conséquence sur la convergence de la méthode du point fixe).

Sous les hypothèses du théorème L.1, on a pour tout $n \in \mathbb{N}^$,*

$$|x_n - l| \leq \frac{k^n}{1-k} |x_1 - x_0| = \frac{k^n}{1-k} |g(x_0) - x_0|. \quad (\text{L.6})$$

DÉMONSTRATION. Voir preuve dans [RDO88, Théorème du point fixe p. 63, section 2.4.3]. Ce théorème de point fixe se démontre dans un cadre beaucoup plus général que le théorème L.1. Il repose sur la majoration suivante : pour tout $n \geq 1$: il existe $\xi \in [x_{n-1}, x_n]$ tel que

$$|x_{n+1} - x_n| = |g(x_n) - g(x_{n-1})| = |g'(\xi)| |x_n - x_{n-1}|,$$

et donc

$$|x_{n+1} - x_n| = |g'(\xi)| |x_n - x_{n-1}|. \quad (\text{L.7})$$

D'après (L.1), il vient donc

$$|x_{n+1} - x_n| \leq k |x_n - x_{n-1}| \quad (\text{L.8})$$

et donc par récurrence sur n

$$|x_{n+1} - x_n| \leq k^n |x_1 - x_0|, \quad (\text{L.9})$$

On considère ensuite deux entiers n et p en supposant par exemple $p < n$. On a par inégalité triangulaire :

$$|x_p - x_n| \leq \sum_{q=0}^{n-p-1} |x_{p+q} - x_{p+q+1}| \quad (\text{L.10})$$

et donc, d'après (L.9),

$$|x_p - x_n| \leq |x_1 - x_0| \sum_{q=0}^{n-p-1} k^{p+q}.$$

En utilisant la somme des termes d'une suite géométrique, il vient compte tenu de $0 \leq k < 1$,

$$|x_p - x_n| \leq \frac{k^p}{1-k} |x_1 - x_0|. \quad (\text{L.11})$$

Cette inégalité assure d'une part que la suite x_n a une limite égale à l , l'unique point fixe de g et d'autre part, en y faisant $n \rightarrow +\infty$:

$$\forall p \geq 1, \quad |x_p - a| \leq \frac{k^p}{1-k} |x_1 - x_0|. \quad (\text{L.12})$$

ce qui est donc (L.6). □

Donnons une dernière majoration, utile parfois en pratique :

LEMME L.3 (Autre conséquence sur la convergence de la méthode du point fixe).

Sous les hypothèses du théorème L.1, on a pour tout $n \in \mathbb{N}^$,*

$$|x_n - l| \leq \frac{k}{1-k} |x_n - x_{n-1}|. \quad (\text{L.13})$$

DÉMONSTRATION. On a effet, d'après (L.1), de façon analogue à (L.7)

$$\begin{aligned} |x_n - l| &= |g(x_{n-1}) - g(l)|, \\ &= |g'(\xi)| |x_{n-1} - l|, \\ &\leq k |x_{n-1} - l|, \\ &\leq k(|x_{n-1} - x_n| + |x_n - l|), \end{aligned}$$

et donc

$$(1-k)|x_n - l| \leq k|x_{n-1} - x_n|,$$

ce qui permet de conclure car $1-k > 0$. □

Ce résultat fournit une preuve alternative du théorème L.2.

AUTRE PREUVE DU THÉORÈME L.2. On réécrit (L.9) sous la forme

$$|x_n - x_{n-1}| \leq k^{n-1} |x_1 - x_0|,$$

ce qui permet de conclure directement en utilisant (L.13). □

Convergence des méthodes d'ordre p

Reprenons les résultats de l'annexe L et adaptons-les à une méthode d'ordre p . Dans toute cette annexe, la suite (x_n) est définie par $x_{n+1} = g(x_n)$ et on suppose que cette suite converge vers l , point fixe de g . On note

$$\forall n \in \mathbb{N}, \quad e_n = x_n - l. \quad (\text{M.1})$$

Rappelons la [BM03, Définition 4.3] : voir la définition A.22.

Rappelons la [BM03, Proposition D.1].

PROPOSITION M.1.

- Si l'ordre de convergence p est égal à un, il existe $\delta \in [0, 1[$ et un entier N tels que

$$\forall n \in \mathbb{N}, \quad |e_{n+N}| \leq \delta^n |e_N|. \quad (\text{M.2})$$

- Si l'ordre de convergence p est strictement supérieur à un, il existe $\delta \in [0, 1[$, $\gamma \in \mathbb{R}_+$ et un entier N tels que

$$\forall n \in \mathbb{N}, \quad |e_{n+N}| \leq \gamma \delta^{(p^n)}. \quad (\text{M.3})$$

DÉMONSTRATION.

- Si l'ordre de convergence est égal à un, il existe $\delta \in [0, 1[$ et un entier N tels que, pour tout entier k :

$$|e_{N+k+1}| \leq \delta |e_{N+k}|.$$

On en déduit

$$|e_{N+n}| \leq \delta^n |e_N|.$$

- De même, si l'ordre de convergence est strictement supérieur à un, il existe $D \in \mathbb{R}_+$ tel que, pour tout entier k et pour tout entier N ,

$$|e_{N+k+1}| \leq D |e_{N+k}|^p.$$

Ainsi, pour $k = 0$, on a

$$|e_{N+1}| \leq D |e_N|^p.$$

Pour $k = 1$, il vient donc

$$|e_{N+2}| \leq D |e_{N+1}|^p \leq D(D |e_N|^p)^p = D^{1+p} |e_N|^{(p^2)}.$$

On montre par une récurrence immédiate sur n que

$$\forall n \in \mathbb{N}, \quad |e_{n+N}| \leq D^{1+p+p^2+\dots+p^{n-1}} |e_N|^{(p^n)},$$

ce qui implique, puisque $p > 1$:

$$|e_{n+N}| \leq D^{\frac{p^n-1}{p-1}} |e_N|^{(p^n)} = D^{\frac{-1}{p-1}} \left(D^{\frac{1}{p-1}} |e_N| \right)^{(p^n)}.$$

La majoration (M.2) est valable puisque la convergence est *a fortiori* d'ordre 1. Ainsi e_n tend vers zéro et en prenant N assez grand

$$D^{\frac{1}{p-1}} |e_N| < 1,$$

ce qui nous permet de conclure. □

REMARQUE M.2. Dans le cas où $p \geq 2$, notons que cette preuve fournit les valeurs de γ et δ données par (avec N assez grand)

$$\gamma = D^{\frac{1}{1-p}}, \quad (\text{M.4a})$$

$$\delta = D^{\frac{1}{p-1}} |e_N| < 1. \quad (\text{M.4b})$$

Si on note $[\alpha, \beta]$ l'intervalle d'étude, si on choisit $N = 0$ et que l'on suppose

$$D^{\frac{1}{p-1}} |\beta - \alpha| < 1, \quad (\text{M.5})$$

alors on a

$$\gamma = D^{\frac{1}{1-p}}, \quad (\text{M.6a})$$

$$\delta = D^{\frac{1}{p-1}} (\beta - \alpha), \quad (\text{M.6b})$$

$$\forall n \in \mathbb{N}, \quad |e_n| \leq \gamma \delta^{(p^n)}. \quad (\text{M.6c})$$

Rappelons le [BM03, Corollaire D.4.].

COROLLAIRE M.3. *Si l'ordre de convergence p est strictement supérieur à un, on a à l'infini*

$$r_{n+1} \sim pr_n, \quad (\text{M.7})$$

ce qui signifie, que pour n assez grand, à chaque étape, le nombre de décimales exactes est asymptotiquement multiplié par l'ordre de convergence de la méthode.

DÉMONSTRATION. Voir [BM03, Corollaire D.4.]. □

Nous concluons en donnant une généralisation du résultat (L.6) pour des méthodes d'ordre p . L'inconvénient de la majoration (M.6c) n'est pertinente d'après (M.6b) que si $\beta - \alpha$ est assez petit pour assurer $D < 1$.

THÉORÈME M.4. *On suppose que la fonction g est de classe C^p sur $[\alpha, \beta]$, qu'elle laisse $I = [\alpha, \beta]$ stable, que sont vérifiées les hypothèses (A.51) et (A.52) avec $p \geq 2$. On considère M telle que*

$$\forall x \in [\alpha, \beta], \quad \left| g^{(p)}(x) \right| \leq M. \quad (\text{M.8})$$

et on suppose que (L.1) a lieu¹. Posons

$$A = \frac{M}{(p-1)!} (\alpha - \beta)^{p-1}, \quad (\text{M.9a})$$

$$B = k^{p-1} \in [0, 1[. \quad (\text{M.9b})$$

On pose, pour tout $n \in \mathbb{N}^*$,

$$F(n) = A^n B^{\frac{(n)(n+1)}{2}} \sum_{q=0}^{\infty} A^q B^{\frac{q^2+2nq+q}{2}}. \quad (\text{M.10})$$

Alors

$$\lim_{n \rightarrow \infty} F(n) = 0 \quad (\text{M.11})$$

et

$$\forall n \in \mathbb{N}^*, \quad |x_n - l| \leq |x_1 - x_0| F(n). \quad (\text{M.12})$$

1. ce qui est vrai d'après (A.51) si $\beta - \alpha$ est assez petit.

DÉMONSTRATION. Pour tout $x \in [\alpha, \beta]$, la formule de Taylor-Young appliquée à g' sur $[l, x]$ implique qu'il existe $\xi \in [l, x]$ tel que

$$g'(x) = g'(l) + (x-l)g''(l) + \dots + \frac{1}{(p-2)!}g^{(p-1)}(l)(x-l)^{p-2} + \frac{1}{(p-1)!}g^p(l)(x-l)^{p-1},$$

soit, compte tenu de (A.51) et de (M.8),

$$\forall x \in [\alpha, \beta] \quad |g'(x)| \leq \frac{M}{(p-1)!} |x-l|^{p-1}, \quad (\text{M.13})$$

Si, de plus $x \in [x_n, x_{n+1}]$ ou $x \in [x_{n+1}, x_n]$, alors

$$|x-l| \leq \max(|x_n-l|, |x_{n+1}-l|),$$

et, d'après (L.3), toujours valable

$$\leq |x_0-l| \max(k^n, k^{n+1}),$$

et, puisque $k < 1$

$$= |x_0-l| k^n.$$

On a donc, d'après (M.13),

$$\text{Si } x \in [x_n, x_{n+1}] \text{ ou } x \in [x_{n+1}, x_n], \quad |g'(x)| \leq \frac{M}{(p-1)!} (k^n |x_0-l|)^{p-1}. \quad (\text{M.14})$$

Par ailleurs, (L.7) est toujours valable et fournit, grâce à (M.14),

$$\begin{aligned} |x_{n+1} - x_n| &= |g'(\xi)| |x_n - x_{n-1}|, \\ &\leq \frac{M}{(p-1)!} (k^n |x_0-l|)^{p-1} |x_n - x_{n-1}|, \\ &\leq \frac{M}{(p-1)!} (k^n |\alpha-\beta|)^{p-1} |x_n - x_{n-1}|. \end{aligned}$$

Bref, en considérant A et B définis par (M.9), on a

$$\forall n \in \mathbb{N}^*, \quad |x_{n+1} - x_n| \leq AB^n |x_n - x_{n-1}|. \quad (\text{M.15})$$

En raisonnant de façon proche de l'établissement par récurrence de (L.3), on a

$$\begin{aligned} |x_{n+1} - x_n| &\leq AB^n |x_n - x_{n-1}|, \\ &\leq A^2 B^{n+(n-1)} |x_{n-1} - x_{n-2}|, \\ &\vdots \\ &\leq A^n B^{n+(n-1)+(n-2)+\dots+1} |x_1 - x_0|, \\ &\leq A^n B^{\frac{n(n+1)}{2}} |x_1 - x_0|, \end{aligned}$$

soit

$$\forall n \in \mathbb{N}^*, \quad |x_{n+1} - x_n| \leq A^n B^{\frac{n(n+1)}{2}} |x_n - x_{n-1}|. \quad (\text{M.16})$$

On considère ensuite deux entiers n et p en supposant par exemple $p < n$. On écrit de nouveau l'inégalité triangulaire (L.10) et il vient

$$\begin{aligned} |x_p - x_n| &\leq \sum_{q=0}^{n-p-1} |x_{p+q} - x_{p+q+1}|, \\ &\leq \sum_{q=0}^{n-p-1} A^{p+q} B^{\frac{(p+q)(p+q+1)}{2}} |x_1 - x_0|, \\ &\leq A^p B^{\frac{p(p+1)}{2}} |x_1 - x_0| \sum_{q=0}^{n-p-1} A^q B^{\frac{(p+q)(p+q+1)}{2} - \frac{p(p+1)}{2}}, \\ &\leq A^p B^{\frac{p(p+1)}{2}} |x_1 - x_0| \sum_{q=0}^{n-p-1} A^q B^{\frac{q^2+2pq+q}{2}}, \end{aligned}$$

Ainsi,

$$\forall n, p \in \mathbb{N}^*, \quad n > p \implies |x_p - x_n| \leq A^p B^{\frac{p(p+1)}{2}} |x_1 - x_0| \sum_{q=0}^{n-p-1} A^q B^{\frac{q^2+2pq+q}{2}}. \quad (\text{M.17})$$

À p fixé, la série de terme général $A^q B^{\frac{q^2+2pq+q}{2}}$ est convergente. En effet, on a

$$A^q B^{\frac{q^2+2pq+q}{2}} \leq A^q B^{\frac{q^2}{2}}$$

et donc

$$q^2 A^q B^{\frac{(q^2+2pq+q)(p+1)}{2}} = \exp\left(\frac{q^2}{2} \ln B + q \ln A + 2 \ln q\right),$$

qui tend vers zéro quand q tend vers l'infini et donc

$$A^q B^{\frac{q^2+2pq+q}{2}} = o\left(\frac{1}{q^2}\right).$$

Dans (M.17), on peut faire tendre n vers l'infini, ce qui donne

$$\forall p \in \mathbb{N}^*, \quad |x_p - l| \leq A^p B^{\frac{p(p+1)}{2}} |x_1 - x_0| \sum_{q=0}^{\infty} A^q B^{\frac{q^2+2pq+q}{2}}.$$

On considère F défini par (M.10). On a donc démontré (M.12). Pour conclure, montrons (M.11). Il suffit d'écrire

$$A^p B^{\frac{p(p+1)}{2}} \sum_{q=0}^{\infty} A^q B^{\frac{q^2+2pq+q}{2}} \leq A^p B^{\frac{p(p+1)}{2}} \sum_{q=0}^{\infty} A^q B^{\frac{q^2}{2}} = 0 \left(A^p B^{\frac{p(p+1)}{2}} \right),$$

en remarquant comme précédemment que la série de terme général $A^q B^{\frac{q^2}{2}}$ est convergent. \square

L'intérêt de ce théorème est qu'elle est valable pour tout A , le nombre B étant toujours plus petit que 1. Sa difficulté réside dans le calcul de la somme (M.10), qui peut s'approcher par troncature et contrôle du reste.

Exemple numérique issu des TD ?

Calcul explicite de la somme (M.10) ?

Compléments sur la divergence de la méthode du point fixe (sous forme d'un exercice corrigé)

Afin d'illustrer les propositions A.16 et A.17, nous donnons dans cette annexe, un exercice donné en examen en Informatique 3A, à l'automne 2018, puis nous concluons par quelques remarques.

Énoncé

Soit la fonction g définie par

$$\forall x \in I = [0, 1], \quad g(x) = \begin{cases} 2x, & \text{si } x \in [0, 1/2], \\ 2(1-x), & \text{si } x \in [1/2, 1]. \end{cases} \quad (\text{N.1})$$

- (1) (a) Montrer que les points fixes de g sont 0 et $2/3$.
 - (b) Montrer que I est g -stable (i.e., si $x \in I$ alors $g(x) \in I$).
 - (c) Que peut-on en déduire sur la suite de la méthode du point fixe définie par $x_{n+1} = g(x_n)$?
 - (d) Quelles sont les seules limites possibles de cette suite?
- (2) (a) Est-ce que l'on peut affirmer que l'hypothèse (A.31b) de la proposition A.19 page 70 est vérifiée? est vérifiée?
 - (b) Peut-on affirmer que la méthode du point fixe converge?
- (3) (a) Est-ce que l'on peut affirmer que l'hypothèse (A.30c) de la proposition A.17 est vérifiée?
 - (b) Peut-on affirmer que la méthode du point fixe ne converge pas?
- (4) (a) Calculer les 6 premières valeurs de la suite x_n du point fixe pour $x_0 = 1/24$. On s'efforcera de calculer les différentes valeurs de x_n sous forme de fraction. On pourra faire un petit graphique. Est-ce que dans ce cas, la suite du point fixe converge?
 - (b) Calculer les 5 premières valeurs de la suite x_n du point fixe pour $x_0 = 1/7$. On s'efforcera de calculer les différentes valeurs de x_n sous forme de fraction. On pourra faire un petit graphique. Est-ce que dans ce cas, la suite du point fixe converge?
 - (c) (i) Calculer les 9 premières valeurs de la suite x_n du point fixe pour $x_0 = 1/4\pi$. Avec du courage, on pourra déterminer x_n sous la forme $a_n + \pi b_n$ où a_n et b_n sont des entiers (pour $n \geq 2$).
 - (ii) Que remarquez-vous?

Corrigé

Pour plus de détails, on pourra consulter [BM03, Exercice 4.9 et TP 4.N] ainsi que https://www.univers-ti-nspire.fr/files/pdf/14-th_point_fixe-TNS21.pdf.

Soit la fonction g définie par

$$\forall x \in I = [0, 1], \quad g(x) = \begin{cases} 2x, & \text{si } x \in [0, 1/2], \\ 2(1-x), & \text{si } x \in [1/2, 1]. \end{cases} \quad (\text{N.2})$$

- (1) (a) Si on résoud $g(x) = x$, on a deux cas :
- Si $x \in [0, 1/2]$, cela est équivalent à $2x = x$ et donc $x = 0$ qui est bien dans $[0, 1/2]$.
 - Si, au contraire, on a $x \in [1/2, 1]$, cela est équivalent à $2(1-x) = x$ et donc $3x = 2$, soit encore $x = 2/3$ qui est bien dans $[0, 1/2]$.

Ainsi,

$$\text{les points fixes de } g \text{ sont } 0 \text{ et } 2/3. \quad (\text{N.3})$$

- (b) Soit $x \in [0, 1]$. On a deux cas :
- Si $x \in [0, 1/2]$, alors $g(x) = 2x \in [0, 2 \times 1/2] = [0, 1]$.
 - Si $x \in [1/2, 1]$, alors $g(x) = 2(1-x) \in [0, 2 \times 1/2] = [0, 1]$.

Ainsi

$$I \text{ est } g\text{-stable}. \quad (\text{N.4})$$

- (c) De cela, on peut déduire par récurrence sur n , que si $x_0 \in I$, alors, pour tout n , x_n est dans I , et c'est tout !
- (d) Puisque g est continue, les seules limites possibles de x_n sont les points fixes, c'est-à-dire 0 et $2/3$.

$$\text{Les seules limites possibles de } x_n \text{ sont } 0 \text{ et } 2/3. \quad (\text{N.5})$$

- (2) (a) g est dérivable sauf en $1/2$. Il est clair que

$$\forall x \in I \setminus \{1/2\}, \quad |g'(x)| = 2, \quad (\text{N.6})$$

et donc l'hypothèse (A.31b) de la proposition A.19 page 70 n'est pas vérifiée.

- (b) Ainsi, les hypothèses de la proposition A.19 ne sont pas assurées. Celle-ci assurait que la convergence du point fixe avait lieu ; si les hypothèses ne sont pas vérifiées, on ne peut donc affirmer que la méthode converge (ou diverge d'ailleurs!).
- (3) (a) D'après (N.6), l'hypothèse (A.30c) de la proposition A.17 page 70 n'est pas vérifiée.
- (b) Comme précédemment, on ne peut pas affirmer que la méthode du point fixe ne converge pas (ou converge!).

REMARQUE N.1. On peut en fait réduire I à $I_1 = [0, \varepsilon]$ ou $I_2 = [2/3 - \varepsilon, 2/3 + \varepsilon]$ (avec $0 < \varepsilon < 1/3$), ce qui rend vraie l'hypothèse (A.30c) de la proposition A.17 page 70. Malheureusement, dans ce cas, l'hypothèse (A.30b) de la proposition A.17 page 70, n'est pas valable. En effet, si par exemple, x n'appartient pas à I_1 , on ne peut affirmer que $g(x)$ (qui existe) n'appartient pas à I_1 . En effet, x peut appartenir à $[1 - \varepsilon/2, 1]$. Dans ce cas, $g(x)$ appartient à $[2(1-1), 2(1-1 + \varepsilon/2)] = [0, \varepsilon] = I_1$. De même, si x n'appartient pas à I_2 , il peut appartenir à $[1/3 - \varepsilon/2, 1/3 + \varepsilon/2]$ et dans ce cas $g(x)$ appartient à $[2(1/3 - \varepsilon/2), 2(1/3 + \varepsilon/2)] = [2/3 - \varepsilon, 2/3 + \varepsilon] = I_2$.

- (4) (a) On obtient les valeurs données dans les tableaux N.1 ou N.2. Voir aussi la figure N.1. On constate qu'une fois les 4 premières valeurs de x_n passées ($\{1/24, 1/12, 1/6, 1/3\}$), la valeur de x_n est constamment égale à $2/3$, point fixe de g .

Dans ce cas-là, la suite converge donc vers $2/3$.

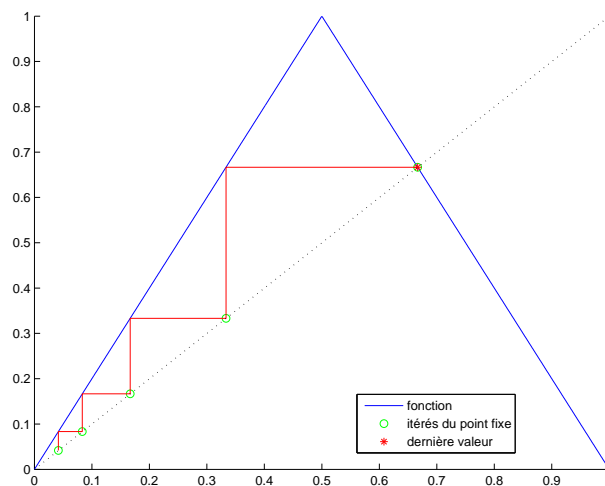
- (b) On obtient les valeurs données dans les tableaux N.3 ou N.4. Voir aussi la figure N.2. On constate qu'une fois la première valeur de x_n passée ($\{1/7\}$), les valeurs de x_n présentent un cycle : $\{2/7, 4/7, 6/7\}$, puisque $g(6/7) = 2/7$.
- Dans ce cas-là, la suite ne converge pas, puisqu'elle "oscille" en permanence.

- (c)
- (i)

n	x_n
0	1/24
1	1/12
2	1/6
3	1/3
4	2/3
5	2/3

TABLE N.1. Valeurs de x_n pour $x_0 = 1/24$.

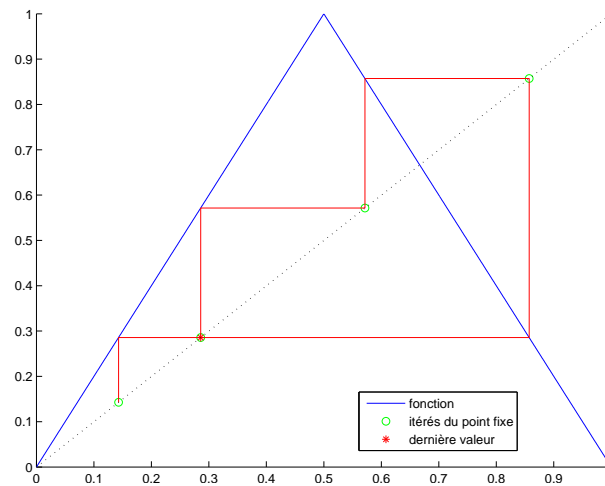
n	x_n
0	0.041666666666667
1	0.083333333333333
2	0.166666666666667
3	0.333333333333333
4	0.666666666666667
5	0.666666666666667

TABLE N.2. Valeurs (numériques) de x_n pour $x_0 = 1/24$.FIGURE N.1. méthode du point fixe pour $x_0 = 1/24$.

n	x_n
0	$1/7$
1	$2/7$
2	$4/7$
3	$6/7$
4	$2/7$

TABLE N.3. Valeurs de x_n pour $x_0 = 1/7$.

n	x_n
0	0.142857142857143
1	0.285714285714286
2	0.571428571428571
3	0.857142857142857
4	0.285714285714286

TABLE N.4. Valeurs (numériques) de x_n pour $x_0 = 1/7$.FIGURE N.2. méthode du point fixe pour $x_0 = 1/7$.

n	a_n	a_n
0	0	1/4
1	2	-1/2
2	4	-1
3	-6	2
4	-12	4
5	26	-8
6	-50	16
7	-100	32
8	202	-64
9	-402	128
10	-804	256
11	-1608	512
12	-3216	1024
13	6434	-2048
14	12868	-4096
15	25736	-8192
16	51472	-16384
17	102944	-32768
18	205888	-65536
19	-411774	131072
20	823550	-262144
21	1647100	-524288
22	-3294198	1048576
23	6588398	-2097152
24	-13176794	4194304
25	26353590	-8388608
26	-52707178	16777216
27	105414358	-33554432
28	-210828714	67108864
29	-421657428	134217728
30	-843314856	268435456
31	1686629714	-536870912
32	-3373259426	1073741824
33	-6746518852	2147483648
34	-13493037704	4294967296
35	26986075410	-8589934592
36	-53972150818	17179869184
37	-107944301636	34359738368
38	-215888603272	68719476736
39	-431777206544	137438953472
40	863554413090	-274877906944
41	-1727108826178	549755813888
42	3454217652358	-1099511627776
43	6908435304716	-2199023255552
44	-13816870609430	4398046511104
45	27633741218862	-8796093022208
46	-55267482437722	17592186044416
47	-110534964875444	35184372088832
48	-221069929750888	70368744177664
49	442139859501778	-140737488355328
50	884279719003556	-281474976710656
51	-1768559438007110	562949953421312
52	-3537118876014220	1125899906842624
53	-7074237752028440	2251799813685248
54	-14148475504056880	4503599627370496
55	28296951008113762	-9007199254740992
56	-56593902016227522	18014398509481984
57	-113187804032455044	36028797018963968
58	-226375608064910088	72057594037927936
59	452751216129820178	-144115188075855872
60	905502432259640356	-288230376151711744

TABLE N.5. Valeurs de $x_n = a_n + \pi b_n$ pour $x_0 = 1/4\pi$.

n	x_n
0	0.785398163397448
1	0.429203673205103
2	0.858407346410207
3	0.283185307179586
4	0.566370614359173
5	0.867258771281654
6	0.265482457436692
7	0.530964914873384
8	0.938070170253233
9	0.123859659493535
10	0.247719318987069
11	0.495438637974138
12	0.990877275948276
13	0.018245448103448
14	0.036490896206895
15	0.072981792413791
16	0.145963584827581
17	0.291927169655162
18	0.583854339310324
19	0.832291321379352
20	0.335417357241296
21	0.670834714482593
22	0.658330571034814
23	0.683338857930372
24	0.633322284139257
25	0.733355431721486
26	0.533289136557027
27	0.933421726885945
28	0.133156546228109
29	0.266313092456218
30	0.532626184912436
31	0.934747630175127
32	0.130504739649746
33	0.261009479299491
34	0.522018958598983
35	0.955962082802035
36	0.088075834395931
37	0.176151668791861
38	0.352303337583722
39	0.704606675167445
40	0.590786649665111
41	0.818426700669779
42	0.363146598660442
43	0.726293197320884
44	0.547413605358232
45	0.905172789283535
46	0.189654421432930
47	0.379308842865860
48	0.758617685731720
49	0.482764628536561
50	0.965529257073122
51	0.068941485853756
52	0.137882971707512
53	0.275765943415025
54	0.551531886830049
55	0.896936226339902
56	0.206127547320196
57	0.412255094640392
58	0.824510189280785
59	0.350979621438431
60	0.701959242876861

TABLE N.6. Valeurs (numériques) de x_n pour $x_0 = 1/7$.

On obtient les 60 premières valeurs données dans les tableaux N.5 ou N.6. Voir aussi la figure

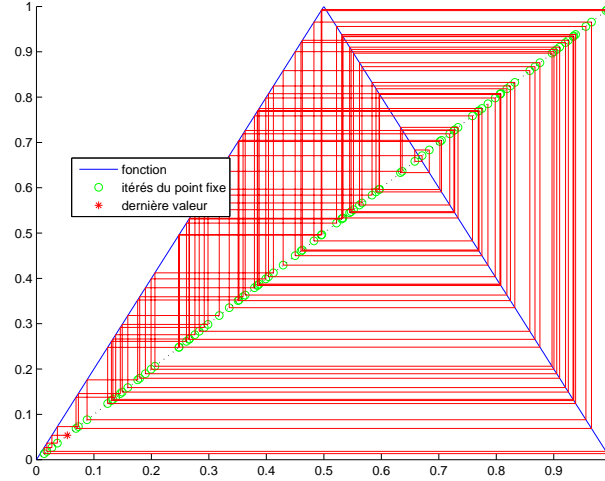


FIGURE N.3. méthode du point fixe pour $x_0 = 1/4 \pi$.

N.3.

On peut montrer que, si $x_0 = 1/4 \pi$, alors

$$\forall n \in \mathbb{N} \setminus \{0, 1\}, \quad \exists!(a_n, b_n) \in \mathbb{Z}, \quad x_n = a_n + \pi b_n. \quad (\text{N.7})$$

En effet, c'est vrai pour $n = 0$ avec $a_0 = 0$ et $b_0 = 1/4$ (qui n'est pas entier!). On démontre ensuite que (N.7) est vrai par récurrence sur n avec a_n et b_n dans \mathbb{Q} . Pour l'existence, on suppose que cela est vrai pour n . Pour $n + 1$, on a, si $x_n \in [0, 1/2]$,

$$x_{n+1} = g(x_n) = 2x_n = 2a_n + \pi(2b_n),$$

et donc

$$\begin{aligned} a_{n+1} &= 2a_n, \\ b_{n+1} &= 2b_n. \end{aligned}$$

Au contraire, si $x_n \in [1/2, 1]$, on a

$$x_{n+1} = g(x_n) = 2(1 - x_n) = 2(1 - a_n) + \pi(-2b_n),$$

et donc

$$\begin{aligned} a_{n+1} &= 2(1 - a_n), \\ b_{n+1} &= -2b_n. \end{aligned}$$

On a donc les formules de récurrence :

$$\forall n \in \mathbb{N}, \quad \begin{cases} a_{n+1} = 2a_n, & b_{n+1} = 2b_n, & \text{si } a_n + \pi b_n \in [0, 1/2], \\ a_{n+1} = 2(1 - a_n), & b_{n+1} = -2b_n, & \text{si } a_n + \pi b_n \in [1/2, 1]. \end{cases} \quad (\text{N.8})$$

Il est clair que, si a_n et b_n sont rationnels, il en est de même pour a_{n+1} et b_{n+1} . L'unicité vient du fait que, si

$$a_n + \pi b_n = a'_n + \pi b'_n,$$

alors

$$(a_n - a'_n) + \pi(b_n - b'_n) = 0,$$

et l'irrationalité de π entraîne la nullité de $a_n - a'_n$ et de $b_n - b'_n$. Enfin, on constate dans le tableau N.5, que, pour $n = 2$, a_n et b_n sont entiers et (N.8) entraîne par récurrence sur n que cela est vrai pour tout $n \geq 2$.

- (ii) On constate que les premières valeurs de x_n sont toutes deux à deux distinctes. On peut montrer qu'il en est de même pour toutes les valeurs de x_n .

REMARQUE N.2. Sur le tableau N.5, on constate que les valeurs de $|a_n|$ et de $|b_n|$ sont de plus en plus grandes. Plus précisément, montrons que

$$\lim_{n \rightarrow +\infty} |a_n| = +\infty, \quad (\text{N.9a})$$

$$\lim_{n \rightarrow +\infty} |b_n| = +\infty. \quad (\text{N.9b})$$

En fait, montrons tout d'abord par récurrence sur n que

$$\forall n \geq 2, \quad |b_n| = 2^{n-2}. \quad (\text{N.10})$$

ce qui implique (N.9b). Dans le tableau N.5, on constate que (N.10) est vrai pour $n = 2$. Supposons maintenant (N.10) vraie pour $n \geq 2$. Montrons-la pour $n + 1$. D'après (N.8) et la récurrence, on a donc

$$|b_{n+1}| = 2|b_n| = 2 \times 2^{n-2} = 2^{n-1},$$

ce qui est bien (N.10) à l'ordre $n + 1$. Montrons maintenant que

$$\forall n \geq 2, \quad |a_n| \geq 2^{n-2}(|a_2| - 2) + 2. \quad (\text{N.11})$$

ce qui implique, d'après la valeur de $|a_2|$ donnée dans le tableau N.5, (N.9a). Montrons d'abord

$$\forall n \geq 2, \quad |a_{n+1}| \geq 2|a_n| - 2. \quad (\text{N.12})$$

On a si $a_n + \pi b_n \in [0, 1/2]$

$$|a_{n+1}| = 2|a_n| \geq 2|a_n| - 2.$$

Si $a_n + \pi b_n \in [1/2, 1]$, on a alors, d'après une inégalité triangulaire,

$$|a_{n+1}| = |2(1 - a_n)| = 2|1 - a_n| = 2|a_n - 1| \geq 2(|a_n| - 1) = 2|a_n| - 2.$$

Concluons enfin en montrant (N.11). On a pour $\alpha = 2$,

$$\alpha = 2\alpha - 2, \quad (\text{N.13})$$

et donc en soustrayant (N.12) à (N.13), on a

$$\forall n \geq 2, \quad |a_{n+1}| - \alpha \geq 2|a_n| - 2\alpha,$$

soit

$$\forall n \geq 2, \quad |a_{n+1}| - \alpha \geq 2(|a_n| - \alpha),$$

et par une récurrence immédiate

$$\forall n \geq 2, \quad |a_n| - \alpha \geq 2^{n-2}(|a_2| - \alpha),$$

et, puisque $\alpha = 2$, (N.11) en découle.

Quelques remarques

- (1) Dans cet exercice, on étudie une méthode de point fixe dont la convergence ne peut être *a priori* affirmée (la proposition A.19 page 70 du cours ne s'applique pas.) et la divergence ne peut être *a priori* affirmée (la proposition A.17 page 70 du cours ne s'applique pas.) En fait, pour tout x_0 appartenant à $I = [0, 1]$, la suite est soit convergente, soit divergente. Plus précisément, on montre que, pour tout $x_0 \in I$, seuls les trois cas exclusifs suivants peuvent se présenter :

- (a) La suite (x_n) possède une infinité de valeurs distinctes ;
 (b) il existe $p \geq 0$ tel que x_0, x_1, \dots, x_p sont deux à deux distincts, puis la suite est stationnaire, c'est-à-dire

$$\forall q \geq p + 1, \quad x_q = x_p. \tag{N.14}$$

- (c) il existe $p \geq 1$ tel que x_0, x_1, \dots, x_p sont deux à deux distincts, puis la suite est cyclique mais non stationnaire, c'est-à-dire : il existe $r \in \{0, \dots, p - 1\}$ tel que

$$x_{p+1} = x_r, \quad x_{p+2} = x_{r+1}, \quad \dots \quad x_{p+k} = x_{p-1-r} \text{ avec } k = p - 1 - r. \tag{N.15}$$

Dans ce cas, on dit que la suite est cyclique d'ordre $k + 1$. Naturellement, les valeurs suivantes des itérés se déduisent par périodicité.

Dans le cas 1a, on peut appliquer alors la proposition A.16 page 70 du cours et il y a divergence. Dans le cas 1c, il y a aussi divergence. Au contraire, dans le cas 1b, il y a convergence. On peut même expliciter la partition U, V et W de I telle que le 1a a lieu si $x_0 \in U$, le 1b a lieu si $x_0 \in V$, le 1c a lieu si $x_0 \in W$. On a : $(\mathbb{R} \setminus \mathbb{Q}) \cup I \subset U$ et en particulier, $\pi/4$ appartient à U . Par exemple, $1/24$ appartient à V . Par exemple, $1/7$ ou $1/84$ appartiennent à W . Les ensembles V et W sont dénombrables. On peut donc affirmer que, pour presque tout $x_0 \in I$, la suite du point fixe est divergente. L'ensemble des x_0 où il y a convergence est cependant dense dans I .

- (2) Sur le plan numérique, les arrondis de calculs rendent les choses plus complexes.

La divergence a en théorie lieu par exemple avec $x_0 = \pi/4$. Si on fait les calculs en symboliques, on obtient bien toutes les valeurs de x_n différentes comme le montre la figure 4(a), ce qui n'est plus vrai pour la figure 4(b). Dans ce dernier cas, pour $n = 51$, on a $x_n = 0$, c'est-à-dire une convergence vers 0.

Si on prend $P = 318310$ points équirépartis dans $[0, 1]$ définis par $x_i = ih$ où i est un entier et $h = \pi/1000000$, on observe dans tous les cas, une convergence vers zéro, alors que, dans ces cas, on devrait observer une divergence !

Si on prend cette fois-ci $P = 1000000$ points aléatoirement répartis dans $[0, 1]$, on observe dans tous les cas, une convergence vers zéro !

- (3) Pour éviter d'avoir ce problème d'arrondis ou d'avoir des temps de calculs trop long dûs aux symbolique, on peut utiliser une autre fonction g définie par

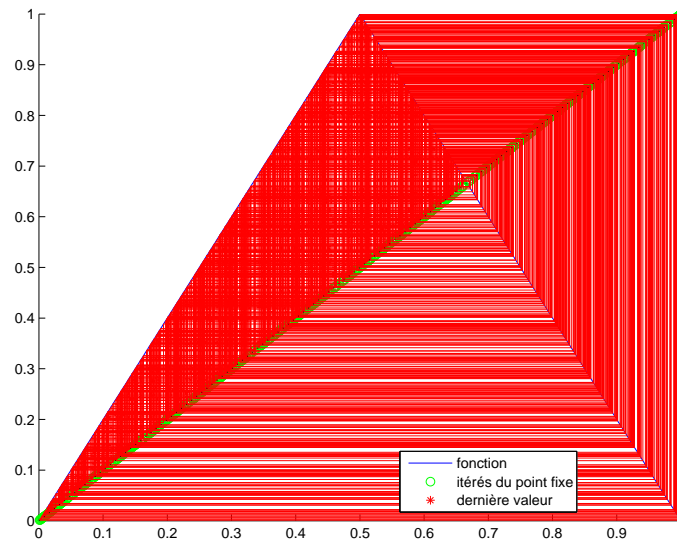
$$\forall x \in [0, 1], \quad g(x) = -4x^2 + 4x. \tag{N.16}$$

Comme précédemment, la fonction g définit une suite (x_n) , qui en général ne converge pas. On peut aussi définir la partition U, V et W de I qui traduit les trois cas précédemment évoqués. Cependant la détermination analytique de U et V n'est pas possible ici. Les calculs se passent beaucoup mieux comme le montre la suite.

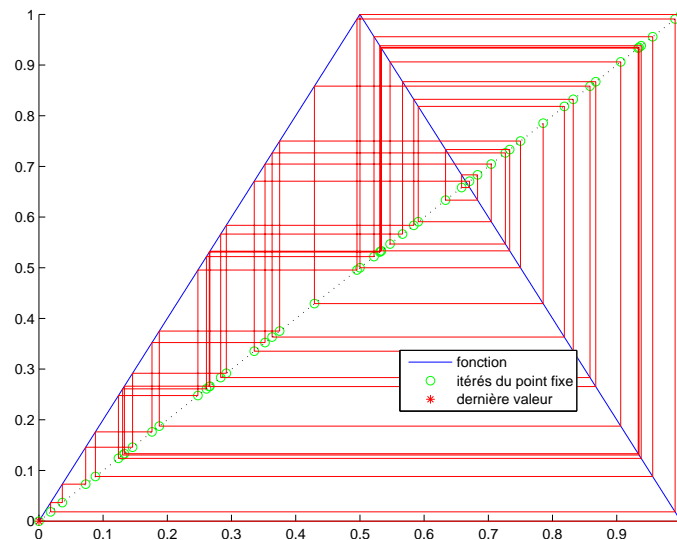
La divergence a de nouveau en théorie lieu par exemple avec $x_0 = \pi/4$, ce qui est bien observé cette fois-ci sur la figure N.5 page 161.

Si on prend $P = 319$ points équirépartis dans $[0, 1]$ définis par $x_i = ih$ où i est un entier et $h = \pi/1000$, on n'observe que dans 0.31 % des cas, une convergence vers zéro, alors que, dans ces cas, on devrait observer une divergence !

Si on prend cette fois-ci $P = 1000$ points aléatoirement répartis dans $[0, 1]$, on observe dans tous les cas, une divergence (aucune valeur égale).



(a) en symbolique



(b) en numérique

FIGURE N.4. calculs pour $x_0 = \pi/4$

- (4) Pour les deux fonctions g définies précédemment, on parle de comportement chaotique de la suite x_n . En effet, elle peut converger ou pas. En cas de divergence, les valeurs qu'elle prend, si elles sont en nombre infini, occupent "tout" $[0, 1]$ de façon imprévisible.

(a)

Plus précisément, on considère les deux valeurs initiales très proches $u_0 = \pi/4$ et $u_0 = \pi/4 + \varepsilon$ où

$$\varepsilon = 1.10^{-12},$$

et on trace les valeurs $|u_n(u_0) - u_n(u_0 + \varepsilon)|$ sur la figure N.6. Cet écart apparaît imprévisible et ne reste jamais très longtemps proche de zéro.

(b)

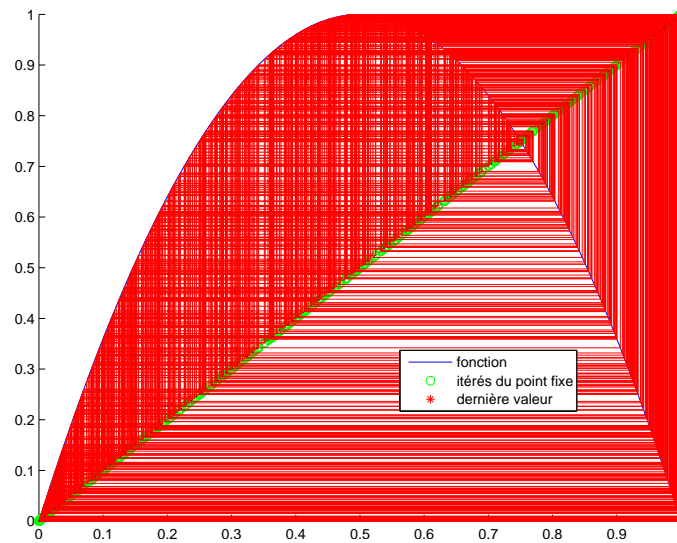


FIGURE N.5. calculs pour $x_0 = \pi/4$ et g définie par (N.16).

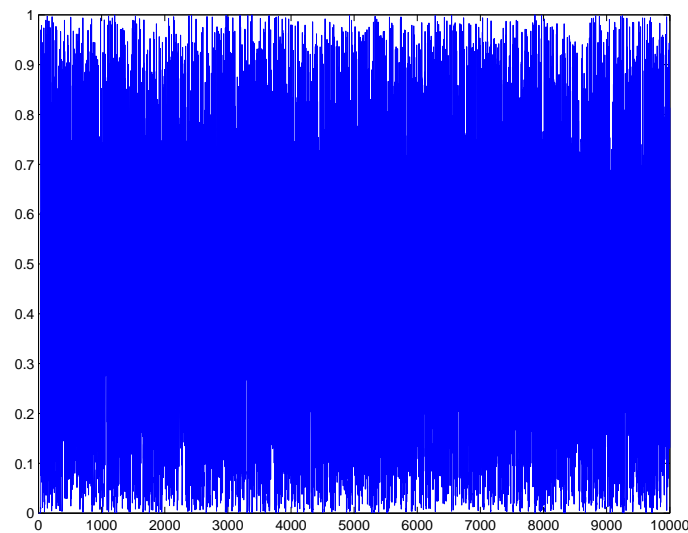


FIGURE N.6. Écart entre x_n pour $x_0 = \pi/4$ et $x_0 = \pi/4 + \varepsilon$ et g définie par (N.16).

Faisons varier n dans $\{1000, 10000, 100000\}$. Par exemple, pour $n = 10000$, les valeurs (toutes différentes) prises par x_n sont représentées en figure N.7. Si on calcule les différentes valeurs de ρ , défini comme le rapport de l'écart maximal divisé par $h = 1/(n + 1)$, on obtient

$$\text{pour } n = 1000, \quad \rho = 10.66,$$

$$\text{pour } n = 10000, \quad \rho = 11.32,$$

$$\text{pour } n = 100000, \quad \rho = 15.41.$$

Cela signifie que, si n augmente, les valeurs que prend x_n , si elles sont en nombre infini, occupent "de plus en plus" $[0, 1]$ et cela, de façon imprévisible, comme le montre par exemple la figure N.8.

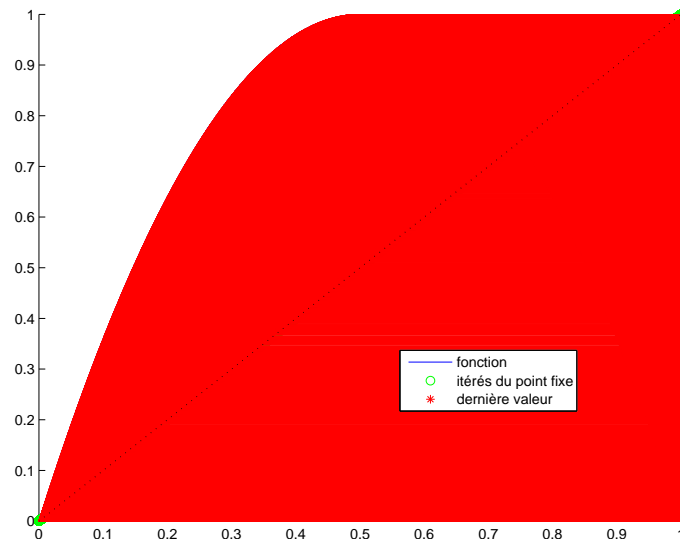


FIGURE N.7. 10000 premières valeurs de x_n pour $x_0 = \pi/4$ et g définie par (N.16).



FIGURE N.8. Zoom des valeurs entre 0.50000 et 0.50154 de x_n pour $x_0 = \pi/4$ et g définie par (N.16).

Bibliographie

- [Bas22] J. BASTIEN. *Mathématiques Fondamentales pour l'Informatique*. Notes de cours de l'UV MFI (Département Informatique) de Polytech Lyon, disponible sur le web : <http://utbmjb.chez-alice.fr/Polytech/index.html>. 2022. 270 pages.
- [BM03] J. BASTIEN et J.-N. MARTIN. *Introduction à l'analyse numérique. Applications sous Matlab*. Ouvrage disponible à la bibliothèque Sciences de Lyon 1 (cote : 519.4 BAS, 4^e étage). Voir <https://www.dunod.com/sciences-techniques/introduction-analyse-numerique-applications-sous-matlab>. Paris : Dunod, 2003. 392 pages.
- [CM84] M. CROUZEIX et A. L. MIGNOT. *Analyse numérique des équations différentielles*. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris, 1984, pages viii+171.
- [DB21] N. DÉBIT et J. BASTIEN. *Méthodes numériques de base*. Notes de cours de l'UV MNB (Département Informatique) de Polytech Lyon, disponible sur le web : <http://utbmjb.chez-alice.fr/Polytech/index.html>. 2021. 288 pages.
- [DB22] N. DÉBIT et J. BASTIEN. *Méthodes numériques de base*. Notes de cours de l'UV MNB (Département Matériaux) de Polytech Lyon, disponible sur le web : <http://utbmjb.chez-alice.fr/Polytech/index.html>. 2022. 326 pages.
- [RDO88] E. RAMIS, C. DESCHAMPS et J. ODOUX. *Cours de mathématiques spéciales. 3. Topologie et éléments d'analyse*. 2^e édition. Ouvrage disponible à la bibliothèque Sciences de Lyon 1 (cote : 510.7 RAM, 4^e étage). Masson, Paris, 1988, pages VIII+362.
- [RDO93] E. RAMIS, C. DESCHAMPS et J. ODOUX. *Cours de mathématiques spéciales. Vol. 1*. Ouvrage disponible à la bibliothèque Sciences de Lyon 1 (cote : 510.7 RAM, 1^{er} étage). Masson, Paris, 1993, pages viii+440.
- [Rey98] A. REY. *Dictionnaire historique de la langue française*. Paris : Dictionnaires Le Robert, 1998.